08-10-00

PCT #

FORM PTO-1390 (REV. 5-93)	U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER
		0342/1D516-US1

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US)

09/601965

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US99/02940	08 February 1999	09 February 1998

TITLE OF INVENTION

A NOVEL FUNGAL PROTEIN CRITICAL FOR EXPRESSION OF FUNGAL PROTEINS

424 Rec'd PCT/PTO 09 AUG 2000

OIPE AUG 09 2000 PATENT AND TRADEMARK OFFICE

APPLICANT(S) FOR DO/EO/US

Craig M. THOMPSON, Fan LONG, and Richard C. WOBBE

Applicant herewith submits to the United States Designated/Elected office (DO/EO/US) the following items and other information:

1. [X] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.

2. [] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S. C. 371.

3. [] This is an express request to begin national examination procedures (35 U.S.C. 371 (f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S. C. 371 (b) and PCT Articles 22 and 39 (1).

4. [X] A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.

5. [X] A copy of the International Application as filed (35 U.S. C. 371 (c) (2))
   a. [] is transmitted herewith (required only if not transmitted by the International Bureau).
   b. [] has been transmitted by the International Bureau
   c. [X] is not required, as the application was filed in the United States Receiving Office (RO/US)

6. [] A translation of the International Application into English (35 U.S. C. 371 (c)2)).

7. [] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371 (c)(3))
   a. [] are transmitted herewith (required only if not transmitted by the International Bureau).
   b. [] have been transmitted by the International Bureau.
   c. [] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [] have not been made and will not be made.

8. [] A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371 (c) (3)).

9. [X] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)) (unsigned).

10. [] A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371 (c)(5)).

Items 11. to 16. below concern other document(s) or information included:

11. [] An Information Disclosure Statement under 37 CFR 1.97 and 1.98 (with references).

12. [] An assignment document for recording. A **separate** cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.

13. [X] A FIRST preliminary amendment.
    [] A SECOND or SUBSEQUENT preliminary amendment.

14. [] A substitute specification.

15. [] A change of power of attorney an/or address letter.

16. [X] Other items or information: **Sequence listing: diskette and paper copy; Statement Pursuant to 37 C.F.R. 1.821; Letter Accompanying U.S. National Phase Filing and copy of the International Preliminary Examination Report**

EXPRESS MAIL CERTIFICATE

8/9/00 EL628221992

Date Label No.

I hereby certify that, on the date indicated above I deposited this paper or fee with the U.S. Postal Service & that it was addressed for delivery to the Commissioner of Patents & Trademarks, Washington D.C. 20231 by "Express Mail Post Office to Addressee" service.

D B Peck D B Peck

Name (Print) Signature

U.S. APPLICATION NO. (if known see 37 C.F.R.1.50)	INTERNATIONAL APPLICATION NO. PCT/GB00/02840	ATTORNEY DOCKET NUMBER
09/601965	534 Rec'd PCT/PTO 09 AUG 2000	0342/1D516-US2

17. [x] The following fees are submitted:	CALCULATIONS	PTO USE ONLY
Basic National Fee (37 CFR 1.492 (a)(1)-(5)): Search Report has been prepared by the EPO [] or JPO [] **$840.00**		
International preliminary examination fee paid to USPTO (37 CFR 1.482) **$670.00**		
No international preliminary examination fee paid to USPTO(37 CFR 4.482) but international search fee paid to USPTO (37 CFR 1.445 (a) (2)... **$690.00**		
Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO..... **$970.00**		
International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4) ...**XXX**..... **$96.00**		
ENTER APPROPRIATE BASIC FEE AMOUNT =	**$96.00**	
Surcharge of $130.00 for furnishing the oath or declaration later than []20 []30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$	

Claims	Number Filed	Number Extra	Rate		
Total Claims	**13-20**	**0**	**0** X $18.00	$	
Independent Claims	**6-3**	**3**	**3** X $78.00	**$234.00**	
Multiple dependent claims(s) (if applicable)		+ 260		$	
			TOTAL OF ABOVE CALCULATIONS =	**$330.00**	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).				$	
			SUBTOTAL =	$	
Processing fee of **$130.00** for furnishing the English translation later the [] 20 [] 30 months from the earliest claimed priority date (37 CFR 1.492(f)).			+	$	
			TOTAL NATIONAL FEE =	**$330.00**	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). the assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property			+	$	
			TOTAL FEES ENCLOSED =	**$330.00**	
				Amount to be refunded	$
				charged.	$

a. [X] A check in the amount of $330.00 to cover the above fees is enclosed.

b. [] Please charge my Deposit Account No.04-0100 in the amount of $ to cover the above fees.

c. [X] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 04-0100. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO
Anne E. Zitron
Darby & Darby P.C.
805 Third Avenue
New York, New York 10022-7513

SIGNATURE Anne E. Zitron

NAME **Anne E. Zitron**

REGISTRATION NO. **41,391**

10-23-00 525 Rec'd PCT/PTO 20 OCT 2000 09/601965

OIPE JC/1/05 OCT 20 2000 PATENT & TRADEMARK OFFICE

Transmittal Letter

Application Number:	09/647683
Filing Date:	September 28, 2000
First Named Inventor:	McHenry
Attorney Docket Number:	1201-009/ddh
Title	*Folding Pocket Knife With Lock*

Enclosed for filing in the above referenced application are the following:

☒ Petition to Make Special (37 C.F.R. § 1.102(d))
☒ Petition Fee of $130.00 (check number 2994)
☒ Other: return receipt postcard

The Commissioner is hereby authorized to charge any additional fees that may be required in connection with prosecution of this application to Deposit Account No. 50-0241

Firm or Individual Name:	Douglas D. Hancock (Registration No. 35,889) **ipsolon llp**
Signature	
Date	**October 20, 2000**

Express Mail No.: EL304655642US

Customer Number: 21,034

09/601965

Express Mail Certificate
Date: 8/9/00 Label: EL 628221992

I hereby certify that on the date indicated above I deposited this paper or fee with the U.S. Postal Service and that it was addressed for delivery to the Commissioner of Patents & Trademarks, Washington, DC 20231 by "Express Mail Post Office to Addressee" service.

D B Peek D B Peek
Name (Print) Signature

534 Rec'd PCT/PTC 09 AUG2000

Docket No.: 0342/1D516US2

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of: Craig M. THOMPSON et al.

Serial No.: TBA Art Unit: TBA
(National Phase of PCT/US99/02940)

Filed: Concurrently Examiner: TBA

For: NOVEL FUNGAL PROTEIN CRITICAL FOR EXPRESSION OF FUNGAL PROTEINS

PRELIMINARY AMENDMENT,
SUBMISSION OF SEQUENCE LISTING
AND
STATEMENT PURSUANT TO 37 C.F.R. § 1.821

Hon. Commissioner of
Patents and Trademarks
Washington, DC 20231

Sir:

Please amend the above-identified patent application as follows:

In the Specification:

After page 45 and before the claims, please insert the attached paper copy of the Sequence Listing.

REMARKS

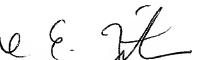
In compliance with 37 C.F.R. § 1.821, applicants herein provide a Paper Copy of a Sequence Listing as well as a Computer Readable Form (diskette).

In addition, applicants herein provide anamendment to the specification directing the entry of the paper copy of the Sequence Listing into the specification after page 45 and before the claims.

Finally, pursuant to Rule 821(f), applicants herein state that the content of the attached paper entitled "SEQUENCE LISTING" and of the accompanying identically labeled diskette, specifically the ASCII-encoded file therein labeled "seqlist.txt", are identical.

Consideration of this Amendment and submission of the enclosed diskette and paper copy are respectfully requested.

Respectfully submitted,



Anne E. Zitron, Ph.D.
Registration No. 41,391
Agent for Applicants

DARBY & DARBY, P.C.
805 Third Avenue
New York, N.Y. 10022
Phone (212) 527-7700

::ODMA\WORLDOX\M:\0342\1D516\AEZ1051.WPD

13 Rec'd PCT/PTO 09 AUG 2000
09/601965

A NOVEL FUNGAL PROTEIN CRITICAL FOR EXPRESSION OF FUNGAL PROTEINS

This application claims priority under 35 U.S.C. § 119 from provisional patent application Serial No. 60/074,100, filed February 9, 1998, the entire disclosure of which is incorporated by reference herein in its entirety.

Field of the Invention

This invention pertains to proteins required for activated transcription in yeast and fungi, nucleic acids encoding these proteins, and methods of using these proteins.

Background of the Invention

Most fungi are opportunistic pathogens, producing serious disease only in compromised individuals. As the result of an aging population and an increase in the number of immunocompromised patients, specifically patients with acquired immunodeficiency syndrome (AIDS), patients undergoing cancer and corticosteroid therapy, as well as in patients undergoing organ transplantation, fungal infections are increasing rapidly.

Most infections begin by colonization of either the skin, a mucosal membrane, or the respiratory epithelium. Passage through the initial surface barrier is accomplished through a mechanical break in the epithelium or enzymatic degradation

or spore dissemination. Most fungi are readily killed by neutrophils and are only opportunists, but some species are resistant to phagocytic killing and may infect otherwise healthy individuals.

Fungi parasitize many different tissues. Superficial fungi cause indolent lesions of the skin. Subcutaneous pathogens cause infection through the skin and spread by subcutaneous or lymphatic routes. Opportunistic fungi such as *Aspergillus* are widespread in the environment and are present in normal flora and cause disease mostly in immunocompromised individuals. Systemic fungi are the most virulent and may cause progressive disease leading to deep seated visceral infections in otherwise healthy individuals (see e.g. *Sherris Medical Microbiology, Third Edition*, Kenneth J. Ryan, ed., Appleton & Lange, Norwalk, CT, 1994).

The major fungal pathogens in North America are *Histoplasma capsulatum*, *Coccidioides immitis*, *Blastomyces dermatitides*, *Cryptococcus neoformans*, *Candida* species and *Aspergillus* species (*Medically Important Fungi, Second Edition*, Davise H. Larone, Ed., American Society for Microbiology, Washington, D.C.). *Histoplasma capsulatum* causes histoplasmosis, which may be chronic or progressive and fatal. It is either a localized or disseminated infection, primarily of the reticuloendothelial system. *Coccidioides immitis* causes coccidioidomycosis, a highly infectious disease that is endemic to the southwestern United States and may be a chronic, sometimes fatal infection involving the skin, bone, joints, lymph nodes, adrenal glands and central nervous system. *Blastomyces dermatitides* causes blastomycosis, a chronic infection characterized by suppurative and granulomatous lesions that begins in the lungs and is disseminated to the skin and bones. *Cryptococcus neoformans* causes cryptococcoses, which may be a chronic infection involving the central nervous system. *Candida albicans* is the most frequent cause of candidiasis, which ranges from an acute to a chronic infection involving any part of the body. *Aspergillus fumigatus* is one of the most frequent causes of aspergillosis, which is an opportunistic infection in immunosuppressed individuals.

Fungi are a distinct class of microorganism, most of which are free-

living. They are eukaryotic organisms containing a nuclear membrane, mitochondria and an endoplasmic reticulum. The cell structure includes a rigid cell wall of mannan, glucan, and chitin and a cytoplasmic membrane with a large percentage of ergosterol. The size and morphology of fungi vary. There are monomorphic yeasts and yeast-like organisms including *Candida*, *Cryptococcus*, and *Saccharomyces*. There are monomorphic molds, such as *Aspergillus* and *Coccidioides*. Some thermally dimorphic fungi, such as *Blastomyces dermatitides* and *Histoplasma capsulatum*, grow either in a yeast or mold phase.

Only a handful of agents are active against fungi. For life threatening disease caused by any of these fungi, amphotericin B is the agent of choice. This drug, however, is associated with numerous severe side effects such as fever, dyspnea and tachycardia, and dosage is limited over the lifetime of the patient because of renal toxicity. An agent frequently used concurrently is flucytosine, a nucleoside analog that cannot be used independent of other agents because of the rapid appearance of resistance. Untoward effects of treatment with flucytosine include leukopenia, thrombocytopenia, rash, nausea, vomiting, diarrhea, and severe enterocolitis.

In conditions where the patient's life is not threatened, ketoconazole can be used as a long-term therapy for blastomycosis, histoplasmosis, or coccidioidomycosis. Fluconazole also has a significant role in the treatment of superficial fungal infections. Both compounds are from the same class, the triazoles, and are cytostatic. The emergence of resistance and hepatic toxicity limit the use of triazoles such as fluconazole and ketoconazole. The newest triazole, itraconazole, has similar pharmacokinetics and spectrum of activity as fluconazole. None of the azoles can be used for life threatening or deep seated fungal infections, but they are effective in reducing colonization of fungi such as Candida and for treating superficial mycoses.

All major antifungal agents attack directly or indirectly a component of the cell wall--ergosterol. Amphotericin B and other polyene macrolides interact with ergosterol in the cell membrane and form pores or channels that increase the

permeability of the membrane. Resistant to amphotericin B in mutant strains is accompanied by decreased concentrations of ergosterol in their cell membranes. Imidazoles and triazoles inhibit sterol 14-α-demethylase, a microsomal cytochrome P_{450}-dependent enzyme system. Imidazoles and triazoles thus impair the biosynthesis of ergosterol for the cytoplasmic membrane and lead to the accumulation of 14-α-methyl sterols, which impair certain membrane-bound enzyme systems (See e.g. *The Pharmacological Basis of Therapeutics, Eighth Edition*, Goodman and Gilman, Pergamon Press, 1990).

Development of an effective method and composition for treatment of fungal infections is a critical goal of the pharmaceutical industry. The pharmaceutical industry has made numerous efforts to identify fungal-specific drugs, with only limited success to date. It would be of great value to identify a new class of antifungal drug that blocks a fungal target other than ergosterol. This target should be fungal-specific and should lead to development of a drug that is effective against the organisms that are resistant to current therapy.

Drug development often relies on the screening of a large number of potential inhibitors before a specific lead compound inhibitor is found. Assays developed for such screens are complex and must mimic the physiological activity of the target protein. Thus, it is critical for the development of these screens to define the proteins involved in the targeted process and to have discovered a means of purifying the necessary components of the assay for use in the assay. In addition, it is useful to have clones for the protein components of the assay to facilitate the production of the components.

Therefore, there is a need in the art to identify one or more fungal constituents, preferably polypeptides, that can serve as useful targets for drug intervention, and for methods and compositions for identifying useful anti-fungal agents and treating fungal infections.

Summary of the Invention

The present invention provides an isolated fungal polypeptide, termed TAF-145, that is necessary for activated transcription of particular genes (i.e., gene-specific transcription) in *Candida albicans*. The invention also includes nucleic acid sequences encoding TAF-145, as well as DNA vectors and transformed cells suitable for recombinant expression of this polypeptide. The DNA sequence of *C. albicans* TAF-145 (SEQ ID NO:1) is set forth in Figures 3A-3G.

In one aspect, the present invention provides methods and compositions for inhibiting gene-specific transcription in *C. albicans*, comprising contacting the cell with an agent that selectively interferes with the transcriptional activation activity of the *Candida* TAF complex, preferably with the activity of TAF-145. In one embodiment, the inhibitory agent is a fragment of TAF-145 that inhibits TBP-TAF interaction.

In another aspect, the invention provides a method for high-throughput screening of large numbers of test compounds to identify an agent useful in the treatment of fungal diseases, specifically those caused by *C. albicans*. The method is carried out by exposing the TAF complex or TAF-145 to TATA Box Binding Protein (TBP) in the presence of at least one test compound, followed by determining that the compound inhibits the binding of TBP to the TAF complex or TAF-145. In another embodiment, candidate antifungal agents are identified as those that bind directly to TAF-145, which are identified using methods such as, e.g., those disclosed in U.S. Patent 5,585,277. In yet another embodiment, candidate antifungal agents are identified as those that inhibit the histone acetyl transferase (HAT) activity of TAF-145.

These and other aspects of the present invention will be apparent to those of ordinary skill in the art in light of the present description, claims and drawings.

Brief Description of the Drawings

Figures 1A-1C together illustrate a protein sequence comparison among TAF-145 derived from *Saccharomyces cerevisiae* (SEQ ID NO:3), *S. pombe* (SEQ ID NO:4), *Drosophila* (SEQ ID NO:5), and humans (SEQ ID NO:6).

Figure 2A is a schematic illustration of the cloning strategy used to obtain the sequence of *C. albicans* TAF-145 (SEQ ID NO:1) using degenerate oligonucleotide PCR.

Figure 2B is a photographic illustration of an agarose gel in which TAF-145-specific PCR products are displayed.

Figures 3A-3G together depict an illustration of the entire nucleotide sequence of the *C. albicans* TAF-145 gene (SEQ ID NO:1) and the predicted amino acid sequence (SEQ ID NO:2).

Figures 4A-4D together illustrate a protein sequence comparison between *C. albicans* (SEQ ID NO:2) and *S. cerevisiae* (SEQ ID NO:3) TAF-145.

Figures 5A-5C together illustrate a protein sequence comparison among the TAF-145 proteins from three yeast species: *C. albicans* (SEQ ID NO:2), *S. cerevisiae* (SEQ ID NO:3), and *S. pombe* (SEQ ID NO:4).

Figures 6A-6C together illustrate a sequence comparison among the TAF-145 homologs from *C. albicans* (SEQ ID NO:2), *S. cerevisiae* (SEQ ID NO:3), *S. pombe* (SEQ ID NO:4), *Drosophila* (SEQ ID NO:5), and human (SEQ ID NO:6).

Figure 7A is an photographic illustration of the results of an experiment in which a *S. cerevisiae* strain temperature sensitive for TAF-145 was transformed with DNA encoding either *C. albicans* TAF-145 or *S. cerevisiae* TAF-145. Only the *Saccharomyces*-derived sequence was able to support growth of the strains under restrictive conditions.

Figure 7B is an photographic illustration of the results of an experiment in which a *S. cerevisiae* strain deleted for TAF-145 was transformed with DNA encoding *C. albicans* TAF-145 or *S. cerevisiae* TAF-145. Only the *Saccharomyces*-derived sequence was able to support growth of strains under restrictive conditions.

Figures 8A and 8B together illustrate a sequence comparison among the histone acetyltransferase (HAT) domains from *C. albicans* (SEQ ID NO:7), *S. cerevisiae* (SEQ ID NO:8), *S. pombe* (SEQ ID NO:9) and human (SEQ ID NO:10).

Figure 9 is a photographic illustration of an SDS-PAGE gel of extracts from uninduced (U) and induced (I) cells following Ni-NTA chromatography.

Figure 10A is an illustration describing the electroelution of TAF proteins from an SDS-PAGE gel.

Figure 10B is a photographic illustration of the purity of the electroeluted proteins as determined by a coomassie stained 12 % SDS-PAGE gel.

Figure 11 is a graph depicting the results of the ELISA for testing the specificty of rabbit sera against the recombinant CaTAF145 HAT domain.

Figure 12 is an photographic illustration of a Coomassie stained 12% SDS-PAGE gel and Western blot analysis to examine the expression of recombinant TAF protein in Baculovirus.

Figure 13A is a schematic illustration of the strategy used in the deletion analysis of *C. albicans* TAF145.

Figure 13B is a photographic illustration of the Southern blot analysis to determine if the deletion was successful.

Detailed Description of the Invention

All patent applications, patents, and literature references cited in this specification are hereby incorporated by reference in their entirety. In case of conflict, the present description, including definitions, will control.

Definitions:

"Basal transcription" refers to transcription activity from an RNA polymerase II-directed promoter in the absence of an upstream transcriptional activator.

"Coactivator activity" refers to the activity that allows an upstream transcription factor such as GAL4 or its derivatives to activate transcription from an

RNA polymerase II-directed promoter in an *in vitro* or *in vivo* reconstituted transcription system. Coactivator activity is further defined as an activity that has no effect on basal transcription.

"TATA-box binding protein" or "TBP" is a major component of eukaryotic transcription factors. In fungi and in higher eukaryotes, TBP is isolated as part of a larger protein complex.

"TATA-box binding protein- associated factors" or "TAFs" as used herein refers to polypeptides or complexes of polypeptides required for "coactivator activity" in fungal RNA polymerase II transcription reactions by virtue of their association with TBP.

"Functional homology" between TAF polypeptides or complexes of polypeptides indicates that one or more biochemical properties specific to fungal TAFs are shared. Examples of such properties are: the ability to specifically modulate the transcription from RNA polymerase II-directed promoters in the presence of an upstream activator protein, the capacity to specifically bind TBP as a multisubunit complex or as a single subunit under conditions as described herein; and the presence of histone acetyl (HAT) transferase activity.

"TAF subunits" refers to individual polypeptides that comprise the TAF complex activity. Such polypeptides are distinguished from any polypeptides previously known to be TBP binding proteins. Fungal TAF subunits may be recombinant or purified from natural sources, and may include structural or functional TAF homologues as defined above.

A "fungal-specific epitope" of a fungal TAF subunit comprises a three-dimensional structural conformation presented by a folded or assembled TAF polypeptide that is not presented by the homologous mammalian sequence.

"Modulating transcription" means altering transcription, and includes increasing or decreasing the rate or level of transcription and changing the responsiveness of transcription to regulatory controls.

An "isolated" polypeptide or nucleic acid is defined as one that is

unaccompanied by at least some of the material with which it is associated in its natural state. Generally, an isolated polypeptide constitutes at least about 1%, preferably at least about 10%, and more preferably at least about 50% by weight of the total protein in a given sample. Included in the polypeptide weight are alternative forms such as differentially glycosylated or phosphorylated or otherwise post-translationally modified forms. An "isolated" nucleic acid sequence is present as other than a naturally occurring chromosome or transcript in its natural state and typically is removed from at least some of the proteins with which it is normally associated on a natural chromosome. A partially pure nucleotide sequence constitutes at least about 5%, preferably at least about 30%, and more preferably at least about 90% by weight of total nucleic acid present in a given fraction.

Also encompassed by the invention are nucleic acids that are hybridizable to, or derived from, the TAF-145 sequences described above. In one embodiment, the invention relates to isolated nucleic acids capable of hybridizing with the TAF-145 sequences above or with their complements under the hybridization conditions defined below.

-- Prehybridization treatment of the support (nitrocellulose filter or nylon membrane), to which is bound the nucleic acid capable of hybridizing with that of *C. albicans* TAF-145, at 65°C for 6 hours with a solution having the following composition: 4 x SSC, 10 x Denhardt (1X Denhardt is 1% Ficoll, 1% polyvinylpyrrolidone, 1% BSA (bovine serum albumin); 1 x SSC consists of 0.15M of NaCl and 0.015M of sodium citrate, pH 7);

-- Replacement of the pre-hybridization solution in contact with the support by a buffer solution having the following composition: 4 x SSC, 1 x Denhardt, 25 mM NaPO4, pH 7, 2 mM EDTA, 0.5% SDS, 100 µg/ml of sonicated salmon sperm DNA containing a nucleic acid derived from the sequence of the TAF-145 as probe, in particular a radioactive probe, and previously denatured by a treatment at 100°C for 3 minutes;

-- Incubation for 12 hours at 65°C;

-- Successive washings with the following solutions: (i) four washings with 2 x SSC, 1 x Denhardt, 0.5% SDS for 45 minutes at 65°C; (ii) two washings with 0.2 x SSC, 0.1 x SSC for 45 minutes at 65°C; and (iii) 0.1 x SSC, 0.1% SDS for 45 minutes at 65°C.

The invention also encompasses any nucleic acid exhibiting the property of hybridizing specifically with the above-described *C. albicans* TAF-145 under the conditions described above, but at 40°C, including successive washings in 2X SSC at 45°C for 15 minutes.

It will be understood that the conditions of hybridization defined above constitute preferred conditions for the hybridization, but are in no way limiting and may be modified without in any way affecting the properties of recognition and hybridization of the probes and nucleic acids mentioned above.

The salt conditions and temperature during the hybridization and the washing of the membranes can be modified in the sense of a greater or lesser stringency without the detection of the hybridization being affected. For example, it is possible to add formamide in order to lower the temperature during hybridization.

Nucleic acids that hybridize to the TAF-145 sequences of the invention may be of any length. In one embodiment, such polynucleotides are at least 25, preferably at least 100 and most preferably at least 200 nucleotides long. In another embodiment, the polynucleotide that hybridizes to the polynucleotide of the invention is of the same length as the polynucleotide of the invention.

TAF-145-Encoding Nucleic Acids and Polypeptides

The present invention encompasses nucleic acid sequences that encode TAF-145 from *C. albicans* and related *Candida* species. Methods for determining the relevant nucleic acid sequences are described in Example 1 below, and the deduced amino acid sequences of a TAF-145 gene, i.e. a gene encoding the 145 kDa TAF polypeptide (TAF-145) isolated from *C. albicans*, is shown in Figure 3. The present invention encompasses DNA and RNA sequences, and sense and antisense sequences.

TAF-encoding sequences according to the present invention may be modified by transitions, transversions, deletions, insertions, or other modifications such as alternative splicing. The invention also encompasses genomic TAF-145 sequences and TAF-145 gene flanking sequences, including TAF-145 regulatory sequences. Nucleic acid sequences encoding TAF-145 polypeptides may also be associated with heterologous sequences, including promoters, enhancers, response elements, signal sequences, polyadenylation sequences, introns, 5'- and 3'- noncoding regions, and the like. Other useful heterologous sequences are known to those skilled in the art. Furthermore, the nucleic acids can be modified to alter stability, solubility, binding affinity and specificity. For example, TAF-145 encoding sequences can be selectively methylated. The nucleic acid sequences of the present invention may also be modified with a label capable of providing a detectable signal, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent molecules, biotin, and the like.

In general, nucleic acid manipulations according to the present invention use methods that are well known in the art, as disclosed in e.g. *Molecular Cloning, A Laboratory Manual* (2nd Ed., Sambrook, Fritsch and Maniatis, Cold Spring Harbor), or *Current Protocols in Molecular Biology* (Eds. Aufubel, Brent, Kingston, More, Feidman, Smith and Stuhl, Greene Publ. Assoc., Wiley-Interscience, NY, NY, 1992).

The *C. albicans* TAF-145 of the present invention has not been disclosed or suggested in the prior art. Although there is some sequence homology between the *C. albicans* (SEQ ID NO:2) and *S. cerevisiae* (SEQ ID NO: 3) TAF-145 molecules in the central region of the protein (48%), there is a much lower sequence identity in the aminoterminal and carboxyterminal domains (19% and 29%, respectively). In addition, there is a 74-amino acid sequence (residues 772 to 845) (SEQ ID NO:11) in *C. albicans* TAF-145, located between the central and C-terminal domains, which is not present in *S. cerevisiae*. Moreover, as shown below, *C. albicans* TAF-145 protein does not complement a *S. cerevisiae* TAF-145 temperature

sensitive mutant protein when cells are grown at the restrictive temperature. Thus, the two molecules differ unpredictably with respect to both structure and function.

The invention also provides vectors comprising nucleic acids encoding *C. albicans* TAF-145 and analogs thereof. A large number of vectors, including plasmid and fungal vectors, have been described for expression in a variety of eukaryotic and prokaryotic hosts. Advantageously, vectors may also include a promotor and/or any other transcriptional regulatory sequence operably linked to the TAF-145 encoding portion. The encoded TAF-145 may be expressed by using any suitable vectors and host cells, using methods disclosed or cited herein or otherwise known to those skilled in the relevant art. The particular choice of vector/host is not critical to the invention.

Vectors will often include one or more replication systems for cloning or expression, one or more markers for selection in the host, e.g. antibiotic resistance, and one or more expression cassettes. The inserted TAF-145 coding sequences may be synthesized, isolated from natural sources, prepared as hybrids, etc. Ligation of the coding sequences to the transcriptional regulatory sequences may be achieved by known methods. Suitable host cells may be transformed/transfected/infected by any suitable method including electroporation, $CaCl_2$ mediated DNA uptake, fungal infection, microinjection, microprojectile, or other established methods.

Appropriate host cells include bacteria, archebacteria, fungi, especially yeast, and plant and animal cells, especially mammalian cells. Of particular interest are *E. coli, B. Subtilis, Saccharomyces cerevisiae, Pichia postoris, Candida albicans*, other *Candida* species such as *C. tropicalis, C. parapsilosis, C. krusei*, and *C. glabrata. Aspergillus* species, *SF9* cells, C129 cells, 293 cells, *Neurospora*, CHO cells, COS cells, HeLa cells, and immortalized mammalian myeloid and lymphoid cell lines. Preferred replication systems include M13, ColE1, SV40, baculovirus, lambda, adenovirus, and the like. A large number of transcription initiation and termination regulatory regions have been isolated and shown to be effective in the transcription and translation of heterologous proteins in the various hosts. Examples

of these regions, methods of isolation, manner of manipulation, etc. are known in the art. Under appropriate expression conditions, host cells can be used as a source of recombinantly produced TAF-145.

Nucleic acids encoding TAF-145 polypeptides may also be introduced into cells by recombination events. For example, such a sequence can be introduced into a cell, and thereby effect homologous recombination at the site of an endogenous gene encoding TAF-145, an analog or pseudogene thereof, or a sequence with substantial identity to a TAF-145-encoding gene. Other recombination-based methods such as nonhomologous recombinations, deletion of endogenous gene by homologous recombination, especially in pluripotent cells, may also be used.

The present invention encompasses TAF complexes and subunits purified from wild-type and genetically altered strains of *C. albicans* or recombinantly produced in a non-native context. Yeast TAF complexes comprise about nine polypeptides, or closely related families of polypeptides. The complexes, and polypeptide components thereof, may be isolated by virtue of their affinity for fungal or human TBP, by the use of chromatographic procedures that take advantage of physico-chemical characteristics of the complexes or of individual subunits, or by binding to TAF-specific antibodies. The isolated complexes may contain all, or only a subset, of the total known complement of TAF subunits. TAF multisubunit complexes may also be reconstituted and purified from translation products of subunit genes, or from recombinantly produced TAF subunits. It is also contemplated that additional TAF subunit polypeptides will be identified using methods disclosed herein, and will be used in practicing the present invention.

In one embodiment, a baculovirus expression system permits the recombinant TAF-145 to be modified, processed and transported within a eukaryotic system. In another embodiment, assembly of the TAF complex, or binding of preassembled TAF complexes to TBP, is performed in a reconstituted cell-free system using partially purified or substantially purified components. For example, TAF complexes, or components thereof, may be adsorbed to the surface of a microtiter

plate, and incubated with radiolabelled TBP protein. Functional binding of TBP to TAF complexes or components will result in the association of detectable radioactivity with the plate.

C. albicans TAF-145 according to the invention may be isolated from wild-type or mutant fungal cells, or from heterologous organisms or cells (including, but not limited to, bacteria, fungi, insect, plant, and mammalian cells) including fungal cells into which a fungal-derived protein-coding sequence has been introduced and expressed. Furthermore, the TAF-145 sequence may be part of recombinant fusion proteins. Alternatively, TAF-145 polypeptides may be chemically synthesized by commercially available automated procedures, including, without limitation, exclusive solid phase synthesis, partial solid phase methods, fragment condensation or classical solution synthesis.

Methods for polypeptide purification are well-known in the art, including, without limitation, preparative disc-gel electrophoresis, isoelectric focusing, HPLC, reversed-phase HPLC, gel filtration, ion exchange and partition chromatography, and countercurrent distribution. For some purposes, it is preferable to produce TAF-145 in a recombinant system in which the fungal protein contains an additional sequence tag that facilitates purification, such as, but not limited to, a polyhistidine sequence. The polypeptide can then be purified from a crude lysate of the host cell by chromatography on an appropriate solid-phase matrix. Alternatively, antibodies produced against TAF-145 or against peptides derived therefrom can be used as purification reagents. Other purification methods are possible.

The present invention also encompasses derivatives and homologues of *C. albicans* TAF-145. For some purposes, nucleic acid sequences encoding the peptides may be altered by substitutions, additions, or deletions that provide for functionally equivalent molecules, i.e., function-conservative variants. For example, one or more amino acid residues within the sequence can be substituted by another amino acid of similar properties, such as, for example, positively charged amino acids (arginine, lysine, and histidine); negatively charged amino acids (aspartate and

glutamate); polar neutral amino acids; and non-polar amino acids.

TAF-145 polypeptides can be modified by methods known in the art. For example, TAF-145 may be phosphorylated or dephosphorylated, glycosylated or deglycosylated, and the like. Especially useful are modifications that alter TAF-145 solubility, membrane transportability, stability, and binding specificity and affinity. Some examples include fatty acid-acylation, proteolysis, and mutations in TBP interaction domains that stabilize binding.

TAF-145 of the present invention may also be modified with a label capable of providing a detectable signal, for example, at a heart muscle kinase labeling site, either directly or indirectly. Exemplary labels include radioisotopes, fluorescent compounds, etc. Such labeled TAFs thereof find use, for example, as probes in expression screening assays for proteins that interact with TAF, or in assays for TAF binding to TBP.

Identification of Functionally Important TAF Domains and Binding Partners

The polypeptides, protein complexes, and nucleic acids sequences of the present invention find use in the discovery, design, and development of pharmaceutically useful antifungal agents. The following embodiments of the present invention are directed towards elucidating epitopes and interactions of TAF-145 that can be selectively interfered with in a therapeutically beneficial manner.

In one embodiment, the known sequence of TAF-145 is used to design synthetic peptides comprising portions of the sequence. These peptides range from about 15 to about 50 amino acids in length. Peptides under 60 amino acids in length may be synthesized routinely using commercially available automated synthesizers. The peptides are then added to a cell-free assembly reaction containing, e.g., immobilized TAF complexes and soluble radiolabelled TBP. Determining which synthetic peptides inhibit some interaction of TAF-145 e.g. with other TAF subunits or with other factors, using routine experimentation, identifies different functional domains or epitopes of TAF-145. For example, a peptide fragment derived from

TAF-145 that is found using the above-described method to inhibit the binding of TBP to TAF-145 or to a TAF complex is likely to represent a region of TAF-145 that interacts directly with TBP. In a similar manner, associational domains of different TAF subunits that are involved in interactions among subunits, or between TAF subunits or complexes and other transcriptional components, may be systematically identified. These peptides may themselves constitute useful therapeutic reagents, or may serve as the basis for design and formulation of pharmacologically active compositions.

In another embodiment, important functional domains of TAFs are identified using classical and reverse genetic methods that are well-known in the art. For example, a nested set of deletion mutants can be prepared from the TAF-145 sequence. In this embodiment, progressively longer amino-terminal and carboxy-terminal deletions can be engineered in the TAF-145 sequence. The resulting set of mutant sequences can be individually expressed in a fungal strain under conditions in which the wild-type version of the TAF is not expressed (see, e.g., Example 1 below, in which *S. cerevisiae* was used as a host). By monitoring the function of each mutant, it is possible to identify different regions of the TAF-145 polypeptide that are critical for function i.e. functional domains or epitopes. Based on such studies, using methods that are well-known in the art, it is possible to selectively introduce defined mutations into different regions of the polypeptide and characterize the variant protein's activity using a similar functional analysis.

An important aspect of the present invention is the selection of functionally important domains or epitopes of *Candida* TAF-145 subunits that are structurally and/or functionally distinct from their mammlian homologues. Such domains are particularly useful as targets for antifungal drugs. In the case of TAF-145, the *Candida* version differs in several important respects from its human homologue, TAF-250. *Candida* TAF-145 is approximately half the size of human TAF-250, and the homologous regions display an amino acid similarity and identity of only 58% and 33%, respectively. *Candida* TAF-145 lacks the carboxy terminal

half of its human counterpart that contains the proposed "Bromo domains" and a region rich in acidic amino acid residues.

Identification of important structural and functional domains of TAFs according to the present invention enables the design and production of useful TAF-derived nucleic acid and peptide-based compounds. For example, fusion proteins may be produced between an important TAF domain and e.g., an enzymatically active fragment of a DNA endonculease. The resulting fusion protein, which can be produced in a fungal cell following introduction into the cell of the hybrid DNA operably linked to an expression vector, finds use in modulating TAF-dependent gene transcription. Other useful TAF fusion partners include sequences useful for immobilization. For example, sequences derived from glutathione-S-transferase (GST) provide a binding site for immobilized glutathione, and sequences that form an epitope recognized by an available monoclonal antibody (e.g., 12CA5 monoclonal antibody) provide a binding site for the immobilized antibody.

In another example, particular serine, threonine, or tyrosine residues in a TAF sequence may be identified as functionally important sites for phosphorylation of TAF. See e.g., methods disclosed in Roberts et al. (1991) *Science* **253**, 1022-1026, and in Wegner et al. (1992) *Science* **256**, 370-373. Phosphorylation of TAF subunits may be involved in modulating the transcription activation activity of Polymerase II transcribed genes. Identification of these residues will enable, first, the radiolabelling of TAF subunits with γ-^{32}P-ATP. Furthermore, if phosphorylation of a particular residue is necessary for transcriptional activity, phosphorylation inhibitors may be designed to block activity.

The nucleic acids encoding TAF-145 may also be used to identify other nuclear factors that interact with TAF-145. In this embodiment, a yeast cDNA library containing fusion genes of cDNA joined with DNA encoding the activation domain of a transcription factor (e.g., Gal4) is co-transfected with fusion genes encoding a portion of TAF and the DNA binding domain of a transcription factor. Clones encoding TAF binding proteins are able to complement the transcription factor

and are identified through transcription of a reporter gene. See, e.g., Fields et al. (1989) *Nature* **340**: 245-246,and Chien et al., (1991) *Proc. Natl. Acad. Sci. USA* **88**:9578-9582. It is contemplated that these additional binding partners for TAF will provide additional targets for antifungal drug therapy.

Anti-TAF Antibodies

The present invention encompasses antibodies that are specific for TAF-145 complexes or subunits identified as described above. The antibodies may be polyclonal or monclonal, and may distinguish TAFs from other nuclear proteins, discriminate TAFs from those derived from different species, identify associational or other functional domains, and the like. Such antibodies are conveniently made using the methods and compositions disclosed in Harlow and Lane, *Antibodies, A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988, other references cited herein, as well as immunological and hybridoma technologies known to those in the art. Where natural or synthetic TAF-derived peptides are used to induce an TAF-specific immune response, the peptides may be conveniently coupled to an suitable carrier such as KLH and administered in a suitable adjuvant such as Freunds. Preferably, selected peptides are coupled to a lysine core carrier substantially according to the methods of Tam (1988) *Proc. Natl. Acad. Sci. USA* **85**:5409-5413. The resulting antibodies may be modified to a monovalent form e.g. Fab, FAB', or FV. Anti-idiotypic antibodies, especially internal imaging anti-idiotypic antibodies, may also be prepared using known methods.

In one embodiment, purified *C. albicans* TAF-145 is used to immunize mice, after which their spleens are removed, and splenocytes used to form cell hybrids with myeloma cells and obtain clones of antibody-secreted cells according to conventional techniques. The resulting monoclonal antibodies are screened using *in vitro* assays such as those described above for the following activities: binding to TAF-145, inhibition of TAF-145 incorporation into multimeric TAF complexes, and inhibition of TAF-145 - TBP interaction.

In another embodiment, the entire TAF complex is used as an immunogen as above, and the resulting monoclonal antibodies are screened for their activity in inhibiting the *in vitro* assembly of any component of the TAF complex.

Anti-TAF antibodies may be used to identify and quantify TAF components, using immunoassays such as ELISA, EMIT, CEDIA, SLIFA, and the like. Anti-TAF antibodies may also be used to block the transcriptional function of, e.g., TAF-145 by inhibiting formation of complexes between TAF subunits or between assembled TAF complexes and other transcription components, or by immunodepleting cell extracts or transcription reactions of TAF components. In addition, these antibodies can be used to identify, isolate, and purify TAFs from different sources, and to perform subcellular and histochemical localization studies.

In one embodiment, polyclonal antibodies against the HAT domain of *C. albicans* TAF145 (aa 339-766) were generated by injection of purified recombinant protein into rabbits (Robert Sargeant, Ramona CA). The sera from these rabbits were then screened by ELISA (Figure 11) for their ability to recognize the recombinant CaTAF145 HAT domain (Example 5)

High-Throughput Drug Screening

The present invention encompasses the identification of agents useful in modulating fungal gene transcription, particularly the transcription of genes by RNA Polymerase II in a TAF-dependent manner. In a preferred embodiment, a high-throughput screening protocol is used to survey a large number of test compounds for their ability to interfere with TAF-dependent processes.

Test inhibitory compounds are screened from large libraries of synthetic or natural compounds. Numerous means are currently used for random and directed synthesis of saccharide, peptide, and nucleic acid based compounds. Synthetic compound libraries are commercially available from Maybridge Chemical Co. (Trevillet, Cornwall, UK), Comgenex (Princeton, NJ), Brandon Associates (Merrimack, NH), and Microsource (New Milford, CT). A rare chemical library is

available from Aldrich (Milwaukee, WI). Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available from e.g. Pan Laboratories (Bothell, WA) or MycoSearch (NC), or are readily produceable. Additionally, natural and synthetically produced libraries and compounds are readily modified through conventional chemical, physical, and biochemical means.

Useful inhibitory agents are identified using any suitable assay that employs TAF-145 or TAF-encoding nucleic acids. As examples, protein binding assays, nucleic acid binding assays and gel shift assays are useful approaches.

In one set of embodiments, TAF complexes or TAF-145 as provided by the present invention are to be used in *in vitro* binding assays with either TBP alone or with a combination or subcombination of TBP and general transcription factors (GTFs).

For example, TAF complexes or TAF-145 may be immobilized on microtiter dishes using methods that are standard in the art. The plates are then exposed to radiolabelled TBP e.g. [^{32}P]-TBP in the absence or presence of candidate compounds. Conversely, TBP may be immobilized, and incubated with radiolabelled TAF-145 or TAF complexes in the absence or presence of candidate compounds. Oligonucleotides comprising TBP target sequences may be used in conjunction with TBP and TAF. Postive "hit" compounds are those that inhibit TAF-TBP interaction. In this case, incubation, washing, and radioactivity detection steps can be automated, allowing the screening of a large number of compounds, preferably at least about 1000 compounds per week.

In another embodiment, test compounds are screened to identify those that inhibit the histone acetyltransferase (HAT) activity of TAF-145 (see, e.g., Examples 2 and 3 below). Positive "hit" compounds are those that cause at least about 25%, preferably at least about 50%, and most preferably at least about 75%, inhibition of incorporation of [3H]-acetate into histones using the reaction conditions described in Example 2 below.

In another embodiment, test compounds are screened to identify those that bind TAF-145, using the high-throughput screening methods described in U.S. Patent Nos. 5,585,277 and 5,679,582, in U.S.S.N. 08/547,889, and in PCT published application PCT/US96/19698. These methods may be used for identifying a ligand that binds the TAF 145 protein. According to these methods, a ligand, or a plurality of ligands for TAF 145 target protein is identified by its ability to influence the extent of folding or the rate of folding or unfolding of the target protein. Experimental conditions are chosen so that the target protein is subjected to unfolding, whether reversible or irreversible. If the test ligand binds to the target protein under these conditions, the relative amount of folded:unfolded target protein or the rate of folding or unfolding of the target protein in the presence of the test ligand will be different, i.e. higher or lower, than that observed in the absence of the test ligand. Thus, the method encompasses incubating TAF 145 in the presence and absence of a test ligand or ligands, under conditions in which (in the absence of ligand) the TAF 145 protein would partially or totally unfold. This is followed by analysis of the absolute or relative amounts of folded vs. unfolded target protein or of the rate of folding or unfolding of the target protein.

An important feature of this method is that it will detect any compound that binds to any sequence or domain of the TAF 145, and not only to sequences or domains that are intimately involved in a biological activity or function. The binding sequence, region, or domain may be present on the surface of the TAF 145 when it is in its folded state, or may be buried in the interior of the protein. Some binding sites may only become accessible to ligand binding when the protein is partially or totally unfolded.

Briefly, to carry out this method, the test ligand or ligands are combined with the TAF 145, and the mixture is maintained under appropriate conditions and for a sufficient time to allow binding of the test ligand. Experimental conditions are determined empirically. When testing test ligands, incubation conditions are chosen so that most ligand:TAF 145 protein interactions would be

expected to proceed to completion. The test ligand is present in molar excess relative to the TAF 145. The target protein can be in a soluble form, or, alternatively, can be bound to a solid phase matrix. The matrix may comprise without limitation beads, membrane filters, plastic surfaces, or other suitable solid supports.

In a preferred embodiment, binding of test ligand or ligands to TAF 145 is detected through the use of proteolysis. This assay is based on the increased susceptibility of unfolded, denatured polypeptides to protease digestion relative to that of folded proteins. In this case, the test ligand-TAF 145 protein combination, and a control combination lacking the test ligand, are treated with one or more proteases that act preferentially upon unfolded target protein. After an appropriate period of incubation, the level of intact i.e. unproteolysed target protein is assessed using one of the methods described below e.g. gel electrophoresis and/or immunoassay.

There are two possible outcomes that indicate that the test ligand has bound the target protein. Either 1) a significantly higher, or 2) a significantly lower, absolute amount of intact or degraded protein may be observed in the presence of ligand than in its absence.

Proteases useful in practicing the present invention include without limitation trypsin, chymotrypsin, V8 protease, elastase, carboxypeptidase, proteinase K, thermolysin, papain and subtilisin (all of which can be obtained from Sigma Chemical Co., St. Louis, MO). The most important criterion in selecting a protease or proteases for use in practicing the present invention is that the protease(s) must be capable of digesting the TAF 145 protein under the chosen incubation conditions, and that this activity be preferentially directed towards the unfolded form of the protein. To avoid "false positive" results caused by test ligands that directly inhibit the protease, more than one protease, particularly proteases with different enzymatic mechanisms of action, can be used simultaneously or in parallel assays. In addition, co-factors that are required for the activity of the protease(s) are provided in excess, to avoid false positive results due to test ligands that may sequester these factors.

In a typical embodiment of this method, purified TAF 145 protein is

first taken up to a final concentration of 1-100 μg/ml in a buffer containing 50 mM Tris-HCl, pH 7.5, 10% DMSO, 50 mM NaCl, 10% glycerol, and 1.0 mM DTT. Proteases, such as, for example, proteinase K or thermolysin (proteases with distinct mechanisms of action), are then added individually to a final concentration of 0.2-10.0 μg/ml. Parallel incubations are performed for different time periods ranging from 5 minutes to one hour, preferably 30 minutes, at 4°C, 15°C, 25°C, and 35°C. Reactions are terminated by addition of an appropriate protease inhibitor, such as, for example, phenylmethylsulfonyl chloride (PMSF) to a final concentration of 1mM (for serine proteases), ethylenediaminotetraacetic acid (EDTA) to a final concentration of 20 mM (for metalloproteases), or iodoacetamide (for cysteine proteases). The amount of intact protein remaining in the reaction mixture at the end of the incubation period may then be assessed by any method, including without limitation polyacrylamide gel electrophoresis, ELISA, or binding to nitrocellulose filters. It will be understood that additional experiments employing a narrower range of temperatures can be performed to establish appropriate conditions. This protocol allows the selection of appropriate conditions (e.g., protease concentration and digestion temperature) that result in digestion of approximately 70% of the target protein within a 30 minute incubation period, indicating that a significant degree of unfolding has occurred.

In another embodiment, the relative amount of folded and unfolded TAF 145 protein in the presence and absence of test ligand is assessed by measuring the relative amount of the protein that binds to an appropriate surface. This method takes advantage of the increased propensity of unfolded proteins to adhere to surfaces, which is due to the increased surface area, and decrease in masking of hydrophobic residues, that results from unfolding. If a test ligand binds the TAF 145 (i.e., is a ligand), it may stabilize the folded form of the target protein and decrease its binding to a solid surface. Alternatively, a ligand may stabilize the unfolded form of the protein and increase its binding to a solid surface.

Surfaces suitable for this purpose include without limitation microtiter plates constructed from a variety of treated or untreated plastics, plates treated for

tissue culture or for high protein binding, nitrocellulose filters and PVDF filters.

In another embodiment, the extent to which folded and unfolded target protein are present in the test combination is assessed through the use of antibodies specific for either the unfolded state or the folded state of the protein i.e. denatured-specific ("DS"), or native-specific ("NS") antibodies, respectively. (Breyer, 1989, *J. Biol. Chem.*, **264 (5)**:13348-13354). Polyclonal or monoclonal antibodies are prepared as described above. The resulting antibodies are screened for preferential binding to the TAF 145 protein in its denatured state. These antibodies are used to screen for inhibitors of these interactions.

In another embodiment, molecular chaperones are used to assess the relative levels of folded and unfolded protein in a test combination. Chaperones encompass known proteins that bind unfolded proteins as part of their normal physiological function. In this embodiment, a test combination containing the test ligand and the TAF 145 is exposed to a solid support e.g. microtiter plate or other suitable surface coated with a molecular chaperone, under conditions appropriate for binding the TAF 145 with its ligand and binding of the molecular chaperone to unfolded target protein. The unfolded target protein in the solution will have a greater tendency to bind to the molecular chaperone-covered surface relative to the ligand-stabilized folded target protein. Thus, the ability of the test ligand to bind target protein can be determined by determining the amount of target protein remaining unbound, or the amount bound to the chaperone-coated surface. Alternatively, a competition assay for binding to molecular chaperones can be utilized.

Once conditions are established for high-throughput screening as described above, the protocol is repeated simultaneously with a large number of test ligands at concentrations ranging from 20 to 200 μM. Observation of at least a two-fold increase or decrease in the extent of digestion of the target protein signifies a "hit" compound, i.e., a ligand that binds the target protein. Preferred conditions are those in which between 0.1% and 1% of test ligands are identified as "hit" compounds using this procedure.

In yet another embodiment, the test and control combinations described above can be contacted with a conformation-sensitive fluorescence probe, i.e., a probe that binds preferentially to the folded, unfolded, or molten globule state of the TAF 145 or whose fluorescence properties are in any way affected by the folding status of the TAF 145 protein. Once a particular test compound has been identified as described above, its activity is then confirmed by adding it to an *in vitro* transcription reaction, and measuring its effect on TAF-mediated activated transcription.

It is also contemplated that a useful agent may interfere with the function of TAF-145 but not inhibit TAF-TBP complex assembly. To screen for such compounds, other functional assays are used, such as, e.g., *in vitro* transcription reactions.

Finally, a test compound identified as described above is tested for two properties: its ability to inhibit fungal growth and its lack of effect on mammalian transcription. Fungal growth is measured by any method well-known in the art e.g. optical density of a liquid culture, or colony formation on agar. The lack of effect of a test compound on mammalian TAF-TBP interaction is tested by replacing yeast components with an analogous human *in vitro* transcription system as in, e.g., Manley et al. (1980) *Proc. Natl. Acad. Sci. USA* 77:3855-3859.

It will be understood that a compound that interferes with any aspect of TAF assembly or function is a likely candidate for an antifungal drug. Thus, in a manner similar to that described above for TBP-TAF-145 binding paradigm, binding assays can be routinely devised that measure the interaction of two or more TAF subunits with each other, or the interaction of one or more TAF subunits with other necessary transcription factors.

According to the present invention, useful agents may be found within numerous chemical classes, though typically they are organic compounds, and preferably small organic compounds. Small organic compounds have a molecular weight of more than 50 yet less than about 2,500 daltons, preferably less than about 750, more preferably less than about 250 daltons. Exemplary classes include

peptides, saccharides, steroids, and the like. The compounds may be modified to enhance efficacy, stability, pharmaceutical compatibility, and the like. Structural identification of an agent may be used to identify, generate, or screen additional agents. For example, where peptide agents are identified, they may be modified in a variety of ways to enhance their stability, such as using an unnatural amino acid, such as a D-amino acid, particularly D-alanine, by functionalizing the amino or carboxyl terminus, e.g., for the amino group, acylation or alkylation, and for the carboxyl group, esterification or amidification, or the like. Other methods of stabilization may include encapsulation, for example, in liposomes, etc.

Therapeutic Applications

For therapeutic uses such as the treatment of fungal infections in mammals, the compositions and agents disclosed herein may be administered by any convenient way, such as, e.g., parenterally, conveniently in a physiologically acceptable carrier, such as, e.g., phosphate buffered saline, saline, deionized water, or the like. Typically, the compositions are added to a retained physiological fluid such as blood or synovial fluid. Alternatively, the compositions may comprise creams, ointments, lotions, or sprays for topical use. Generally, the amount administered will be empirically determined, typically in the range of about 10 to 1000 pg/kg of the recipient. For peptide agents, the concentration of will generally be in the range of about 100 to 500 μg/ml in the dose administered. Other additives may be included, such as stabilizers, bactericides, etc. These additives will be present in conventional amounts.

Example 1: Cloning and Characterization of *Candida albicans* TAF-145

A. *Degenerate oligonucleotide PCR*

A protein sequence alignment of the *Saccharomyces cerevisiae* (SEQ ID NO: 3) and TAF-145 counterparts from *Schizosaccharomyces pombe* (SEQ ID NO: 4), *Drosophila* (SEQ ID NO:5) . and human (SEQ ID NO: 6) is shown in Figures 1A-1C.

Three highly conserved regions are boxed. Based on the sequence of these highly conserved regions, degenerate oligonucleotides (see Table 1 below) were designed and used to amplify fragments of DNA from *Candida albicans* strain SC5314 genomic DNA by polymerase chain reaction (PCR) (Figs. 2A, 2B and SEQ ID NO:12 - SEQ ID NO:15).

PCR reactions were performed with Taq DNA polymerase (Promega) in 25 µl of buffer (provided by manufacturer) supplemented with 2 mM $MgCl_2$ and 200 µM dNTP for a total of 30 cycles. Primers (shown in Table 1 below) were used at a concentration of 1.0 µM with 100 ng of DNA and cycling was at 94°C (45 sec), 37°C (1.0 minute), and 72°C (1 minute) for four cycles followed by 94°C (45 sec), 40°C (1.0 minute), and 72°C (1 minute) for 26 cycles.

TABLE 1: Primers

Oligonucleotide	Sequence	Direction	Corresponding *S. cerevisiae* amino acids	SEQ ID NO:
TAF145p1F	5'-CC(A/T)GG(A/T)CC(A/T)AA(C/T)TCIA(A/G)(A/G)-3'	Forward	668-673	12
TAF145p2F	5'-GA(C/T)CC(A/T)AC(T/C/A)GG(A/T)TGTGG(A/T)GAAGG-3'	Forward	817-824	13
TAF145p3R	5'-CCTTTC(A/T)CCACA(A/T)CCAGT(A/T)GG(A/G)TC-3'	Reverse	817-824	14
TAF145p4R	5'-TT(A/G)TT(A/C/T)CA(C/T)CTIA(G/A)TG(T/A)CC-3'	Reverse	1045-1051	15

A major 700 bp fragment of DNA was amplified by use of the TAF145p1F and TAF145p3R primers. The amplified DNA was subcloned into the SrfI site of pCR-Script Amp SK(+) plasmid (Stratagene) and subjected to DNA sequencing (see below).

B. *Cloning and sequence analysis of C. albicans TAF145*

[^{32}P]-labeled DNA was generated by random hexamer priming (Boerhinger Mannheim) of the cloned 700 bp *C. albicans* TAF-145 PCR product. The radiolabeled DNA was used to screen a *C. albicans* genomic DNA cosmid library by colony hybridization (Sambrook et al., 1989). The *C. albicans* genomic DNA

cosmid library was constructed using conventional procedures: the genomic DNA used to construct the library was purified from *C. albicans* strain SC5314 (Fonzi and Irwin, 1993) as described by Philippsen et al. (1991).

A 6.3 kb fragment of DNA containing the full length *C. albicans* TAF-145 (CaTAF145) gene was subcloned from a cosmid clone into pCT538, creating pFL207 (see Table 2 below). All DNA manipulations were performed according to Sambrook et al. (1989). Figures 3A-3G display the entire nucleotide sequence of the *C. albicans* TAF145 gene (SEQ ID NO:1) and the predicted amino acid sequence (SEQ ID NO: 2) . *C. albicans* translates the standard leucine CUG codon as serine (Santos et al., 1996). Amino acids 980 and 1085 are shown as serines to reflect this divergence from the standard genetic code.

Figures 4A-4D compare the amino amino acid sequences of *C. albicans* (SEQ ID NO:2) and *S. cerevisiae* (SEQ ID NO: 3) TAF145. While there is some sequence identity (48%) between *C. albicans* and *S. cerevisiae* in the middle portion of the protein, there is much lower sequence identity in the N-terminus and C-terminus portions (19% and 29%, respectively). The middle portion of the protein contains a putative histone acetyltransferase domain (Mizzen et al., 1996). An additional difference between these two protein sequences is a 74-amino acid sequence (residues 772-845) in *C. albicans* TAF-145 (SEQ ID NO:11) , located between the central and C-terminus domains, which is not present in *S. cerevisiae* TAF-145. Figures 5A-5C show a protein sequence alignment of the TAF-145 proteins from three yeast species: *C. albicans* (SEQ ID NO: 2) , *S. cerevisiae* (SEQ ID NO: 3) , and *S. pombe* (SEQ ID NO: 4) , and Figures 6A-6C display an alignment of the TAF-145-related proteins from *C. albicans* (SEQ ID NO:2) , *S. cerevisiae* (SEQ ID NO: 3) , *S. pombe* (SEQ ID NO: 4) , *Drosophila* (SEQ ID NO: 5) , and human (SEQ ID NO: 6) . All sequence alignments were done with the Lasergene software package from DNASTAR Inc. The alignments illustrate the divergence of the aminoterminal and carboxyterminal domains of these proteins relative to the central domain.

C. *Functional analysis of Candida albicans TAF-145*

Two experiments were performed to determine if a genomic clone of *C. albicans* TAF-145 (CaTAF145) could substitute functionally for *S. cerevisiae* TAF-145 (ScTAF145). Plasmids and yeast strains used in these experiments are described in Tables 2 and 3, respectively. Yeast media, plasmids, and strains were prepared as described (Guthrie and Fink, 1991). Yeast transformations were done using a lithium acetate procedure (Schiestl and Gietz, 1989). Plasmid shuffle techniques were performed as described (Boeke et al., 1987) using 5-fluoro-orotic acid (5-FOA) as a selective agent against URA3 plasmids. All DNA manipulations used to generate the necessary plasmids were performed according to Sambrook et al. (1989).

TABLE 2

Plasmid	Description
pgTAF145	URA3 CEN ScTAF145
pTAF145^{ts-2}	HIS3 CEN ScTAF145^{ts-2}
pCT538	LEU2 2μ
pFL112	LEU2 2μ ScTAF145
pFL207	LEU2 2μ CaTAF145
pCT3	URA3 CEN
pFL103	URA3 CEN ScTAF145
pFL115	URA3 CEN CaTAF145

TABLE 3

S. cerevisiae Strain	Genotype
YSW85	Mata lys2-801 ade2-101 leu2-Δ1 his3-Δ200 ura3-Δ99 taf145::LEU2 [pgTAF145 (URA3 CEN ScTAF145)]
YSW101	Mata ade2-101 leu2-3,112 his3-11 ura3-1 trp1-1 taf145::hisG [pgTAF145 (URA3 CEN ScTAF145)]

CTY507	Mata lys2-801 ade2-101 leu2-Δ1 his3-Δ200 ura3-Δ99 taf145::LEU2 [pTAF145^{ts-2} (HIS3 CEN ScTAF145^{ts-2})]
YFL021	Mata ade2-101 leu2-3,112 his3-11 ura3-1 trp1-1 taf145::hisG [pgTAF145 (URA3 CEN ScTAF145), pFL207 (LEU2 2μ CaTAF145)]
YFL022	Mata ade2-101 leu2-3,112 his3-11 ura3-1 trp1-1 taf145::hisG [pgTAF145 (URA3 CEN ScTAF145), pFL112 (LEU2 2μ ScTAF145)]
YFL023	Mata ade2-101 leu2-3,112 his3-11 ura3-1 trp1-1 taf145::hisG [pgTAF145 (URA3 CEN ScTAF145), pCT538 (LEU2 2μ)]
YFL024	Mata lys2-801 ade2-101 leu2-Δ1 his3-Δ200 ura3-Δ99 taf145::LEU2 [pTAF145^{ts-2} (HIS3 CEN ScTAF145^{ts-2}), pFL115 (URA3 CEN CaTAF145)]
YFL025	Mata lys2-801 ade2-101 leu2-Δ1 his3-Δ200 ura3-Δ99 taf145::LEU2 [pTAF145^{ts-2} (HIS3 CEN ScTAF145^{ts-2}), pFL103 (URA3 CEN ScTAF145)]
YFL026	Mata lys2-801 ade2-101 leu2-Δ1 his3-Δ200 ura3-Δ99 taf145::LEU2 [pTAF145^{ts-2} (HIS3 CEN ScTAF145^{ts-2}), pCT3 (URA3 CEN)]

One experiment took advantage of a temperature-sensitive (ts) mutant. Plasmid shuffle techniques were used to create CTY507 (a *S. cerevisiae* strain with a temperature-sensitive allele of ScTAF145) from YSW85 and pTAF145^{ts-2} (Reese et al., 1994; Walker et al., 1996). CTY507 was then transformed with pCT3 ("empty" vector), pFL103 (ScTAF145), and pFL115 (CaTAF145) to generate *S. cerevisiae* strains YFL026, YFL025, and YFL024, respectively. Agar plates inoculated with YFL024, YFL025, and YFL026 were placed at permissive (30°C) and restrictive (37°C) temperatures. At the permissive temperature, all three strains grew equally well (Figure 7A). In contrast, at the restrictive temperature, only cells transformed with wild-type *S. cerevisiae* TAF145 (YFL025) grew. Cells containing a genomic

clone of *C. albicans* TAF145 (YFL024) were unable to grow at the restrictive temperature.

A second experiment used plasmid shuffle techniques to determine if *C. albicans* TAF-145 could function in place of *S. cerevisiae* TAF-145. *S. cerevisiae* strain YSW101, containing a wild-type version of ScTAF145 on a URA3 plasmid, was transformed with pCT538 (empty vector), pFL112 (ScTAF145), or pFL207 (CaTAF145) to generate *S. cerevisiae* strains YFL023, YFL022, and YFL021, respectively. Cells were placed on agar lacking leucine and selected for *Candida* TAF-145-encoding plasmids either in the absence or presence of 5-FOA (Figure 7B). All three strains grew equally well in the absence of 5-FOA. When the URA3-containing ScTAF145 plasmid was selected against by the presence of 5-FOA, only those cells (YFL022) transformed by ScTAF145 were able to grow. In the absence of a functional *S. cerevisiae* TAF-145, cells containing a genomic clone of *C. albicans* TAF-145 were unable to grow.

Conditional mutant analysis and plasmid shuffle techniques in *S. cerevisiae* show that a genomic clone of *C. albicans* TAF-145 was unable to substitute functionally *in vivo* for *S. cerevisiae* TAF145. In principle, the lack of complementation may be due to a lack of expression of mRNA and/or protein, or to a substitution of serines 980 and 1085 for leucine. A more likely explanation is that one or more of the protein surfaces required for interactions between yeast TAF-145 and its partners may not be conserved between *S. cerevisiae* and *C. albicans*. While the middle portion of the two proteins share significant sequence similarity, the aminoterminal and carboxyterminal domains have diverged considerably (Figures 4A-4D).

D. *Expression and purification of recombinant proteins*

It has recently been reported that the *S. cerevisiae* TAF-145 protein and the human and *Drosophila* counterparts have histone acetyltransferase (HAT) activity associated with them (Mizzen et al., 1996). This activity has been mapped to the middle, most conserved, portion of ScTAF145 (amino acids 354-817) (SEQ ID

NO:8). Figures 8A and 8B display an alignment of the TAF-145 HAT domain from *C. albicans* (SEQ ID NO:7) , *S. cerevisiae* (SEQ ID NO:8) , *S. pombe* (SEQ ID NO: 9) , and human (SEQ ID NO: 10) . The HAT domains derived from the three fungal species share approximately 50% amino acid sequence identity, while there is less than 25% amino acid sequence identity between the fungal and human TAF-145 HAT domains.

The portion of the gene encoding the putative HAT domains from *S. cerevisiae*, *C. albicans*, and human were PCR amplified and subcloned into pET19b (Novagen) to create pFL105, pFL107, and pFL108, respectively. Oligonucleotide sequences (SEQ ID NO:16 - SEQ ID NO:21) are listed in Table 4 below. Plasmids were transformed into *E. coli* BL21(DE3) for protein expression. Upon induction with IPTG, an amino-terminal His tagged protein was expressed.

TABLE 4

Oligo-nucleotide	Organism	Sequence	Direction	Corresponding amino acids and restriction sites created	SEQ ID NO:
pScHAT1	*S. cerevisiae*	5'-CCGCTCGAGAT GACACCCAACTTAAA GTTC-3'	Forward	amino acid 354 to.... XhoI	16
pScHAT3	*S. cerevisiae*	5'-CGCGGATCCAGA GATTTTAGCTTAGAA TC-3'	Reverse	amino acid 787 BamHI	17
pCaHAT3	*C. albicans*	5'-GGAATTC CATATGCTTTTGCTC AACAATCCCTTGGAC -3'	Forward	amino acid 339 to... NdeI	18
pCaHAT4	*C. albicans*	5'-CGCGGATCCCT GCTCTGCTCACCGAA TAACAC-3'	Reverse	amino acid 766 BamHI	19
pHuHAT 6	Human	5'-GGAATTC CATATGAGCCTGGCA GGCTGGCTTCC TTCT-3'	Forward	amino acid 432 to... NdeI	20
pHuHAT 7	Human	5'-CCGCTCGAGTTC TGGAGCAAAAAAGG ATTTCTC-3'	Reverse	amino acid 912 XhoI	21

Restriction enzyme sites are underlined.

His-tagged HAT domain fusion proteins were purified on a nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography matrix (Qiagen) according to manufacturers instructions. Cells were lysed and proteins solubilized under denaturing conditions. Proteins were refolded on the Ni-NTA column and eluted with 0.5 M imidazole. Figure 9 shows the protein profile of extracts from uninduced (U) and induced (I) cells following Ni-NTA chromatography. Addition of IPTG induced proteins of the expected size for the *S. cerevisiae* (Sc) and *C. albicans* (Ca) proteins (Figure 9).

Example 2: Histone acetyltransferase assays

Two assays for histone acetyltransferase (HAT) activity are described below. The first assay is an activity gel assay that permits direct association of polypeptides with acetyltransferase activity (Brownell and Allis, 1995). Purified HAT proteins are analyzed for HAT activity following electrophoresis in SDS/polyacrylamide gels containing calf thymus histones or bovine serum albumin. Samples are dissolved in SDS/PAGE sample buffer, but not boiled, and then loaded onto standard SDS/polyacrylamide gels which have been modified so that protein substrates are dissolved in the resolving gel at 1 mg/ml prior to polymerization. Following electrophoresis, gels are washed for 1 hour at room temperature in buffer A (50 mM Tris-HCl pH8.0, 20%(v/v) isopropanol, 1 mM DTT, 1 mM PMSF, 0.1 mM EDTA). Gels are then incubated in Buffer A containing 8 M urea for 1 hour and then overnight at 4°C in buffer A containing 0.04% Tween 40 (Sigma). The gels are then washed in buffer B (50 mM Tris-HCl pH8.0, 10%(v/v) glycerol, 1 mM DTT, 1 mM PMSF, 0.1 mM EDTA) prior to incubation with [^{3}H]acetyl-CoA (10 μCi) for 30 minutes at 30°C. Finally, the labeled gels are washed with 5% trichloroacetic acid to remove unbound radiolabel and fluorographed.

The second assay is a standard solution assay (Brownell and Allis, 1995). Enzyme samples are incubated at 30°C in a total volume of 50 μl of buffer B (50 mM Tris-HCl pH8.0, 10%(v/v) glycerol, 1 mM DTT, 1 mM PMSF, 0.1 mM

EDTA) and 25 mg of calf thymus histones (Sigma). The reactions are initiated by the addition of [^{3}H]acetyl-CoA (100 nCi; 6.1 Ci/mmol; ICN) to a final concentration of 0.5 μM and are terminated after 10 minutes by spotting the entire mixture onto Whatman P81 filters. [^{3}H]-Acetate incorporation is determined by liquid scintillation and nonspecific counts are subtracted. This assay can be adapted for a high throughput screen for the identification of compounds that inhibit the activity of fungal TAF145 acetyltransferase activity. A counterscreen with the human HAT domain protein is used to identify those inhibitors specific for the fungal proteins.

References for Examples 1 and 2

Bai, Y., Perez, G.M., Beechem, J.M., and Weil, P.A. (1997). Structure-function analysis of *TAF130*: Identification and characterization of a high-affinity TATA-bind ng protein interaction domain in the N terminus of yeast $TAF_{II}130$. Mol. Cell. Biol. *17*, 3081-3093.

Boeke, J., Truehart, J., Natsoulis, B., and Fink, G.R. (1987). 5-Flouro-orotic acid as a selective agent in yeast molecular genetics. Meth. Enzymol. *154*, 164-175.

Brownell, J.E. and Allis, C.D. (1995). An activity gel assay detects a single, catalytically active histone acetyltransferase subunit in *Tetrahymena* macronuclei. Proc. Natl. Acad. Sci. USA *92*, 6364-6368.

Chen, J.-L., Attardi, L.D., Verrijzer, C.P., Yokomori, K., and Tjian, R. (1994). Assembly of recombinant TFIID reveals differential coactivator reguirements for distinct transcriptional activators. Cell *79*, 93-105.

Fonzi, W.A., and Irwin, M.Y. (1993). Isogenic strain construction and gene mapping in *Candida albicans*. Genetics *134*, 717-728.

Guthrie, C., and Fink, G.R. eds. (1991). Guide to Yeast Genetics and Molecular Biology. Academic Press, New York.

Mizzen, C.A. et al. (1996). The $TAF_{II}250$ subunit of TFIID has histone acetyltransferase activity. Cell *87*, 1261-1270.

Nishikawa, J.-I., Kokubo, T., Horikoshi, M., Roeder, R.G., and Nakatani, Y. (1997). *Drosophila* $TAF_{II}230$ and the transcriptional activator VP16 bind competitively to the TATA box-binding domain of the TATA box-binding protein. Proc. Natl. Acad. Sci.

USA *94*, 85-90.

Philippsen, P., Stotz, A., and Scherf, C. (1991). DNA of *Saccharomyces cerevisiae*. Meth. Enzymol. *194*, 169-182.

Reese, J.C., Apone, L., Walker, S.S., Griffin, L.A., and Green, M.R. (1994). Yeast $TAF_{II}s$ in a multisubunit complex required for activated transcription. Nature *371*, 523-527.

Ruppert, S., and Tjian, R. (1995). Human $TAF_{II}250$ interacts with RAP74: implication for RNA polymerase II initiation. Genes Dev. *9*, 2747-2755.

Ruppert, S., Wang, E.H., and Tjian, R. (1993). Cloning and expression of human $TAF_{II}250$: a TBP-associated factor implicated in cell-cycle regulation. Nature *362*, 175-179.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). Molecular Cloning: A Laboratory Manual, Second Edition (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press).

Santos, M.A.S., Perreau, V.M., and Tuite, M.F. (1996). Transfer RNA structural change is a key element in the reassignment of the CUG codon in *Candida albicans*. EMBO *15*, 5060-5068.

Schiestl, R.H. and Gietz, R.D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Gent. *16*, 339-346.

Walker, S.S., Reese, J.C., Apone, L.M., and Green, M.R. (1996). Transcription activation in cells lacking $TAF_{II}s$. Nature *383*, 185-188.

Weinzierl, R.O.J., Dynlacht, B.D., and Tjian, R. (1993). Largest subunit of *Drosophila* transcription factor IID directs assembly of a complex containing TBP and a coactivator. Nature *362*, 511-517.

Example 3: High-throughput Screening of Anti-TAF compounds

A. *TAF binding*

Corning ELISA strip wells (8 wells per strip) are coated with avidin (1.0 µg per well) by incubating avidin (200 µl of a 5 µg/ml stock) in coupling buffer (per liter: 1.6g Na_2CO_3, 2.9 g, $NaHCO_3$, 0.9 g NaN_3) in the well for 12 h at 4°C. The

buffer is decanted, and nonspecific binding sites on the wells are blocked with 1% skim milk in phosphate-buffered saline (PBS) for 1 h at 37°C. Blocking buffer is discarded, and a yeast Pol II promoter-containing oligonucleotide (1 pmol/well) is added to the wells and incubated for 30 minutes at room temperature. The oligonucleotide is double-stranded and contains a biotin tag on the sense strand.

The oligonucleotide-containing solution is then removed, and the wells are washed with 1% milk in PBS. Yeast TBP is mixed with partially purified TAF that had been metabolically labelled with ^{35}S-methionine or purified TAF-145 similarly labelled, all in HEG buffer (0.1 M KCl, 25 mM HEPES pH 7.9, 0.5 mM EDTA, 20% glycerol, 0.01% LDAO, 0.1 M AEBSF, 0.1 M Na metabisulfite, 10 mM β-mercaptoethanol) plus 200 ug/ml bovine serum albumin (BSA).

The protein mixture is then added to the prepared wells and incubated for 30 minutes at room temperature. Samples are then removed, and the wells are washed three times with the PBS/milk solution. Wells are separated and put into scintillation vials, scintillation cocktail is added, and samples are counted in a liquid scintillation counter.

Binding of yeast TAF to the wells is found to be dependent on the presence of TBP, bound in turn to the Pol II promoter-containing oligonucleotide. Small molecules, whether purified or present in natural or synthetic mixtures, are introduced into the assay at concentrations ranging from about 20 to about 200 μM, and appropriate solvent controls are also performed. Compounds that inhibit binding of TAF by more than about 30% are identified, and the inhibitory activity purified if not already available in pure form.

Compounds identified as described above are then tested for their ability to inhibit TBP-dependent transcription in a mammalian cell-free system. Alternatively, human TBP is used in place of yeast TBP in the above-described assay for TAF binding to TBP. Thus, a subset of active compounds are identified that selectively interfere with yeast, and not mammalian, TAF-145 function, i.e., that are fungal-specific.

B. *HAT activity*

To each well of a microtiter dish is added 50 µl of a solution containing purified *C. albicans* TAF-145, buffer B (50 mM Tris-HCl pH8.0, 10%(v/v) glycerol, 1 mM DTT, 1 mM PMSF, 0.1 mM EDTA) and 25 µg of calf thymus histones (Sigma). Small molecules, whether purified or present in natural or synthetic mixtures, are introduced into the assay at concentrations ranging from about 20 to about 200 µM, and appropriate solvent controls are also performed.

The reactions are initiated by the addition of [^{3}H]acetyl-CoA (100 nCi; 6.1 Ci/mmol; ICN) to a final concentration of 0.5 µM and are terminated after 10 minutes by spotting the entire mixture onto Whatman P81 filters. [^{3}H]-Acetate incorporation is determined by liquid scintillation and nonspecific counts are subtracted. Compounds that inhibit [^{3}H]-Acetate incorporation by at least about 25% are identified.

Compounds identified as described above are then tested for their ability to inhibit TBP-dependent transcription in a mammalian cell-free system. Alternatively, human TAF-145 is used in place of *Candida* TAF-145 in the above-described assay for HAT activity. Thus, a subset of active compounds are identified that selectively interfere with yeast, and not mammalian, TAF-145 function, i.e., that are fungal-specific.

Example 4: TAF Protein expression and purification for antibody production

Recombinant clones encoding for the HAT domain of the TAF protein, pFL 107 (*C. albicans* TAF 145 aa 339-766 in pET19b) [Figs. 8A, 8B and SEQ ID NO:7], pFL 110 (*S. cerevisiae* TAF 145 aa 354-835 in pET23a) [Figs. 8A, 8B and SEQ ID NO:8] and, pFl 123 (human TAF 250 aa 433-974 in pET23a) [Figs. 8A, 8B and SEQ ID NO:10] were each seeded into 10 ml of Luria Bertani (LB) broth supplemented with ampicillin (100 µg/ml) and grown overnight shaking at 37°C. The overnight cultures were seeded into 500 ml of LB/amp and grown shaking at 37°C. to mid-late log phase (O.D. of 0.6-0.8 at 600 nm). The cells were induced by the

addition of 0.4 M isopropyl $_{\beta\text{-D}}$-thiogalactopyranoside (IPTG) (Sigma Chemical Co. St. Louis MO) and allowed to express the TAF proteins for 2 hours. The cells were harvested by centrifugation at 10.000 x g at 4°C. for 20 minutes and resuspended in Tris buffer (50mM Tris-HCL, pH. 8.0). The cells were lysed by sonication 3 x 3 minutes on ice using a Branson sonifier 250 (Branson Ultrasonics Corp., Danbury CT) and the soluble portion was separated from the insoluble portion by centrifugation. The insoluble portion was solubilized by adding 3 ml of urea buffer (8 M urea) and incubated for 30 minutes to 2 hours at 37°C. The solubilized proteins were separated from cellular debris by certification 10,000 x g for 20 minutes. SDS-PAGE sample buffer (0.5 M 2-mercapto-ethanol, 10% [wt./vol.] Sodium dodecyl sulfate (SDS), 50% Glycerol, 0.5% Bromophenol blue) was added to the extracts and the samples boiled for 5 minutes.

The solubilized proteins from *C.albicans* (Ca), *S. cerevisiae* (Sc) and the human (Hu) TAF proteins were loaded onto a 12% SDS-polyacrylamide gel (3mm in thickness). The samples were electrophoresed until the bromphenol blue dye front ran to the bottom of the gels. The proteins were visualized by staining the gels with Coomassie blue stain (10% acetic acid, .25% Coomassie Brilliant Blue) and destained overnight with destain (7% acetic acid, 30% methanol). Destain was

09874935 .102000

removed from the gels and the gels were washed 3 x for 10 minutes in deionized water.

For purification of the *C. albicans*, *S. cerevisiae,* and the human TAF proteins. The protein bands were cut from the gel and electroeluted using a Centrilutor device (Amicon Inc., Beverly MA) according to the manufacturer's instructions. The purity of the electroeluted proteins was determined on a coomassie stained 12 % SDS-PAGE gel (Bio-Rad, Hercules CA) (Figure 10B). Each TAF protein band was discrete and free of other contaminating proteins.

The purified protein against the *C. albicans* TAF145 aa 339-766 [SEQ ID NO:7] was then injected into rabbits to produce polyclonal antibodies (see Example 5).

Example 5: Production of antibodies against *C. albicans* TAF145 HAT domain

Polyclonal antibodies directed against a portion of the HAT domain of *C. albicans* TAF145 (aa 339-766) [SEQ ID NO: 7] were generated by injection of purified recombinant protein into rabbits (Robert Sargeant, Ramona CA). Sera from these rabbits were tested for their ability to recognize recombinant CaTAF145 HAT domain in an ELISA assay according to the following method:

1. Varying amounts of CaTAF145 HAT domain were aliquotted in 50 mM Borate, pH8.5, were added to Immulon 4 plates, were incubated for 25°C for 1 hour, and the plates were then washed.

2. The plates were then blocked with 200 μl of a 5% milk solution in a Tris buffer (TBST)

TBST = 10 mM Tris, pH 7.5
150 mM NaCl
0.5% Tween-20

The plates were incubated for 30 minutes at 25°C, and were then washed.

3. The washed plates were then incubated with rabbit antisera at multiple dilutions (1:500, 1:1000, 1:2000, 1:4000) in TBST(5% milk) for 1 hour at 25°C, and were then washed.

4. The plates were further incubated with a 1:1000 dilution of APGAR (Alkaline Phosphatase Conjugated Goat-anti-Rabbit polyclonal antibody Calbiochem) in TBST/5% milk. The secondary antibody was incubated for 1 hour at 25°C, and the plates were again washed.

5. The plates were then incubated with 2 mg/ml PNPP (para-nitro-phenylphosphate) in a 1X diethanolamine buffer (Pierce) for approximately 20 minutes. Once color developed, the plates were read at OD 405 nm

The ELISA assay data for the four dilutions tested is graphically represented in Figure 11.

Example 6: Expression of Human and C. albicans TAF proteins in Baculovirus

The full-length *C. albicans* TAF145 gene and a portion of the human TAF250 gene (aa 1-1218) [SEQ ID NO: 22] were PCR amplified and subcloned into pFastBac1 (Life Technologies Inc., Gaithersburg MD) to create pFL139 (*C. albicans)* and pFL149 (human), respectively. Oligonucleotide sequences (SEQ ID NO: 23- SEQ ID NO:28) are listed in the Table 5 below.

TABLE 5

Oligonucleotide	Organism	Sequence	Direction	Corresponding amino acids and restriction sites created	SEQ ID NO:
pCaHAT15	*C. albicans*	ACGCGTCGACATGC ATCATCATCATCAT CATATGGAGGATCT ACCCAGGGAT	Forward	amino acid 1 to ... 6xHis Tag and *Sal*I site.	23
pCaHAT12	*C. albicans*	ATAGTTAGCGGCCG CACACTGCTGGTGT CAACCAACAA	Reverse	amino acid 1161 NotI	24
pHuHAT11	Human	ACGCGTCGACATGC ATCATCATCATCAT CATATGGGACCCGG CTGCGATTTG	Forward	amino acid 1 to... 6xHis Tag and *Sal*I site.	25
pHuHAT12	Human	GTTGCTCTGCAGCT ATCATGCTATAATA AGC	Reverse	amino acid 894 PstI	26
pHuHAT13	Human	TGATAGCTGCAGAG CAACGACTGAAGG ATGC	Forward	amino acid 895 PstI	27
pHuHAT15	Human	CCGGTACCTTCCCG ATGTTGTTCATCAA AAAG	Reverse	amino acid 1218 KpnI	28

All oligonucleotides are written in the 5' to 3' direction.

Restriction enzyme sites are underlined.

pFL149 was created in two steps: the 2.7 kb PCR product created by oligonucleotides pHuHAT11 and pHuHAT12 was inserted into pFastBac1 to create pFL141 followed by insertion of the 0.9 kb PCR product created by oligonucleotides pHuHAT13 and pHuHAT15.Bacmid DNA was generated by transforming pFL139 and pFL149 into E. coli DH10Bac and selecting on Luria Agar plates containing 50 mg/ml kanamycin, 7 mg/ml gentamicin, 10 mg/ml tetracycline, 100 mg/ml X-gal and, 40 mg/ml IPTG. White colonies were selected for the isolation of recombinant bacmid DNA. Bacmid DNA was isolated according to the BAC-To-BAC Baculovirus Expression Systems Instruction Manual (Life Technologies Inc., Gaithersburg MD).

The preparation of bacmid DNA and transfection of SF-9 cells was done using the Bac-to-Bac™ (Life Technologies Inc., Gaithersburg MD) expression system according to the instruction manual. 5 µl of bacmid mini-prep DNA

(recombinant and wild type) and 6 µl of Celllfectin (Life Technologies Inc., Gaithersburg MD) reagent were each diluted into 100 µl of grace's insect medium supplemented with L-glutamine (Invitrogen Corp., San Diego CA). Each solution was mixed by gently inverting several times. The bacmid DNA and cellfectin solutions were combined and allowed to incubate for 30 minutes at rt. 0.8 ml of grace's minimal media was added to the complex and gently mixed by inverting.

$9x10^5$ sf9 cells were seeded into each well (35mm) of a 6 well plate and allowed to attach for 1 hour at room temperature. The cells were washed 2 x with 2 ml of Grace's minimal media. After the final wash the cells were transfected with bacmid DNA-Cellfectin complex for 6 hours rocking at rt. The trasfection media was removed and 2 ml of Ex-cell™ 420 serum-free insect media (JRH Biosciences, Lenexa KS) was added to each well and incubated for 72 hours at 27°C. After 72 hours the transfected cells were removed from the plates and harvested by low speed centrifugation. The supernatant fluid containing baculovirus was removed and stored at 4°C.

1 x 10^6 High-Five cells, available from Invitrogen, Carlsbad, CA, were seeded in each well of a 6 well plate and allowed to attach for 1 hour at room temperature. The media was removed and the cells infected with 100 µl of a 1:10 dilution of transfection baculovirus supernatant. The cells were rocked for 1 hour at rt. The infected cells were supplemented with 2 ml of Ex-Cell media and incubated for 72 hours at 27°C. After 72 hours the cells were collected by washing the wells several times with media and the cells pelleted by low speed centrifugation. The supernatant media was removed and the cells resuspended in phosphate-buffered saline (PBS), pH 7.5.

Coomassie stain and Western blot analysis was used to examine the expression of recombinant TAF protein in Baculovirus (Figure 12). 5 µl of 5x SDS-PAGE sample buffer was added to 10 µl of the cell suspensions and the samples boiled for 5 minutes. Cellular debris was removed by centrifugation and the proteins (15 µl) separated on 12% SDS-PAGE gels (Bio-Rad, Hercules CA). The gels were

placed in a mini Trans-Blot apparatus (Bio-Rad, Hercules CA) and the proteins were electrophoretically transferred to nitrocellulose membranes. The membranes were blocked for 10 minutes with blocking buffer (5% skim milk, 20 mM Tris, 0.2 M NaCl, 0.1% Tween-20, pH 7.5). Rabbit polyclonal anti-C. albicans TAF 145 antibodies were used to probe the electroblotted infected high-five cell extracts. C. albicans TAF 145 [SEQ ID NO: 2] and Human TAF 250 (aa 1-1218) [SEQ ID NO: 22] proteins were detected using goat anti-rabbit antibodies conjugated with HRP and ECL chemiluminescence (Amersham Pharmacia Biotech, Buckinghamshire England). Western Blot analysis illustrated that the Human (Figure 12; Lane 1) and the *C. albicans* (Figure 12; Lane 2) HAT domains of the TAF protein were expressed in the Baculovirus system and the proteins were of the predicted size, 145 kDa.

Example 7: Deletion analysis of *C. albicans* TAF145

Deletion analysis of *C. albicans* TAF145 was used to establish if CaTAF145 is essential for cell viability. In wild-type *C. albicans,* CaTAF145 is encoded by two alleles (Figure 13B; lane 1).

Homologous recombination was used to integrate two plasmids into the yeast genome. For a schematic diagram illustrating the homologous recombination see Figure 13A. Each of the plasmids contain TAF145 interrupted by selectable markers as follows: ADE (pSIK11) and URA3 (pFL125).

Construction of plasmids pSIK11 and pFL125

pSIK11 was constructed in the following manner:

1) 460 bp CaTAF145 upstream region was PCR amplified with primers TAF145p32f (ACGCGTCGACATCCAAGTTCAAGTTGTCTG) [SEQ ID NO:29] and TAF145p33r (CGCGGATCCGCGCTGCAGTTTTCACATCTTCTT CT TCTGCCA) [SEQ ID NO:30];

2) 518 bp CaTAF145 downstream region was PCR amplified with primers TAF145p34f (AAAACTGCAGCGCGGATCC GCGT GCA GGTGA CGTTATTGGA) [SEQ ID NO:31] and TAF145p35r (ATAGTTTAGCGGCCGCC TTGTGACAA GAAGTGACAC) [SEQ ID NO:32];

3) PCR products from steps 1 and 2 were used as templates and primers TAF145p32f and TAF145p35r were used to generate a single piece of DNA by PCR that contains the upstream and downstream region of CaTAF145;

4) the PCR product from step 3 was cloned into SalI-NotI sites of pBlueScript (Stratagene, La Jolla, CA);

5) ADE2 gene was PCR amplified and inserted into PstI-BamHI sites of the plasmid created in step 5.

pFL125 was constructed similarly to pSIK11 except that in step 5, hisG CaURA3 hisG (Fonzi and Irwin, 1993) was inserted into PstI-BamHI sites.

Yeast medium was prepared as described (Guthrie and Fink, 1991). Transformation of *C. albicans* was performed using a lithium acetate procedure (Schiestl and Gietz, 1989). To create a single disruption of TAF145 in *C. albicans* strain CAI8 (Fonzi and Irwin, 1993), cells were transformed with plasmid pSIK11 (ΔCataf145::ADE2) digested with SalI/NotI and ADE+ prototrophs were selected. Integration of the ΔCataf145::ADE2 cassette at TAF145 was verified by Southern Blotting and using ^{32}P- labeled pSIK11 as a probe (Figure 13B; Lane 2). These singly disrupted strains (CSIK1) were selected for further analysis and then an attempt was made to delete the second TAF145 allele.

The singly disrupted strain CSIK1 (TAF145/ΔCataf145::ADE2) was transformed with pFL125 (ΔCataf145::URA3) digested with SalI/NotI and Ade+ Ura+ transformants were selected. As a control, to determine if pF125

(ΔCataf145::URA3) was capable of integrating into the yeast genome, CAI8 was transformed with SalI/NotI digested pFL125 (ΔCataf145::URA3) and Ura+ transformants were selected. Analysis of a Ura+transformant is shown in Figure 13B, lane 3.

Ura+ Ade+ transformants were screened by Southern blot for the presence of both integrated cassettes (Figure 13B; Lanes 4-10). In 20 out of 20 transformants, the wildtype TAF145 band was still present, indicating that it was not possible to select for transformants with a deletion in the second copy of the TAF145 allele and that TAF145 is essential for cell viability. Figure 13B displays the results of Southern blot analysis for 7 of the 20 transformants. Each transformant selected contained a deletion in only one copy of the TAF145 allele and only one selectable marker, either Ura or Ade but not both. This indicates that TAF145 is essential for cell viability.

.References for Example 7

Fonzi, W.A., and Irwin, M.Y. (1993). Isogenic strain construction and gene mapping in *Candida albicans*. Genetics *134*, 717-728.

Guthrie, C., and Fink, G.R. eds. (1991). Guide to Yeast Genetics and Molecular Biology. Academic Press, New York.

Schiestl, R.H. and Gietz, R.D. (1989). High efficiency transformation of intact yeast cells using single stranded nucleic acids as a carrier. Curr. Genet. *16*, 339-346.

****

Many variations of the present invention will suggest themselves to those skilled in the art in light of the above detailed disclosure. All such modifications are within the full extended scope of the appended claims. All patents and references mentioned in this application are hereby incorporated by reference in their entirety.

WHAT IS CLAIMED IS:

1. An isolated nucleic acid having the sequence depicted in Figure 3, SEQ ID NO: 1.

2. An isolated nucleic acid that hybridize to a nucleic acid as defined in claim 1 under stringent hybridization conditions.

3. A nucleic acid vector comprising a nucleic acid as defined in claim 1 operably linked to a transcription regulatory element.

4. A cell comprising a vector as defined in claim 3.

5. A cell as defined in claim 4, wherein said cell is a member selected from the group consisting of bacterial, fungal, insect, and mammalian cells.

6. A method for producing a polypeptide, which comprises:

(i) culturing a cell as defined in claim 5 under conditions suitable for the expression of *C. albicans* TAF 145 polypeptide; and

(ii) recovering said polypeptide from said culture.

7. An isolated polypeptide having the amino acid sequence depicted in Figure 3, SEQ ID NO:2.

8. A fragment of a polypeptide as defined in claim 7 which inhibits the interaction of said polypeptide and TATA-box Binding Protein (TBP).

9. A method for inhibiting fungal transcription in a *C. albicans* cell comprising contacting said cell with an agent that selectively interferes with the

interaction of the polypeptide of claim 7 and TATA Binding Protein (TBP).

10. An antibody that specifically recognizes *C. albicans* TAF 145 polypeptide.

11. A method for rapid, large-scale screening to identify a ligand that binds to a *C. albicans* TAF 145 protein from a plurality of test compounds not known to bind *C. albicans* TAF 145 protein, said *C. albicans* TAF 145 protein having been incubated in the presence and absence of a plurality of test compounds under conditions sufficient to unfold *C. albicans* TAF 145 protein in the absence of test compounds, which comprises:

detecting an increase or a decrease in the amount of *C. albicans* TAF 145 protein in the folded state; and

determining that said test compound is a ligand that binds to said *C. albicans* TAF 145 protein if there is detected an increase or a decrease in the amount of said *C. albicans* TAF 145 protein in the folded state in the presence of said test compound.

12. A method for rapid, large-scale screening to identify a ligand that binds to a *C. albicans* TAF 145 protein from a plurality of test ligands not known to bind to said *C. albicans* TAF 145 protein, said *C. albicans* TAF 145 protein having been incubated in the presence and absence of a plurality of test ligands which comprises:

subjecting said *C. albicans* TAF 145 protein to unfolding conditions;

detecting an increase or a decrease in the amount of *C. albicans* TAF 145 protein in the folded state; and

determining that said test compound is a ligand that binds to said *C. albicans* TAF 145 protein if there is detected an increase or a decrease in the amount of said *C. albicans* TAF 145 protein in the folded state in the presence of said test compound.

13. A method for rapid, large-scale screening to identify a ligand that binds to a *C. albicans* TAF 145 protein from a plurality of test ligands not known to bind to a *C. albicans* TAF 145 protein which comprises:

incubating said *C. albicans* TAF 145 protein in the presence and absence of a plurality of test ligands;

subjecting said *C. albicans* TAF 145 protein to unfolding conditions; and

detecting an increase or a decrease in the amount of *C. albicans* TAF 145 protein in the folded state, wherein said increase or said decrease identifies a test ligand that binds to said *C. albicans* TAF 145 protein.

SUBSTITUTE SHEET (RULE 26)

FIG. 1A

FIG. 1B

FIG. 1C

DOCKET-2961960

FIG. 2A

S. cerevisiae	P G P N S R	D P T G C G E G . . . . . G H I R T N K
S. pombe	P G P H S R	D P T G I G E G . . . . . G H M K T N K
Drosophila	P G P N S K	D P T G C G E G . . . . . G H M R T N K
Human	P G P N S K	D P T G C G E G . . . . . G H M R T N K

TAF145p1F 5' CC(A/T)GG(A/T)CC(A/T)AA(C/T)TCIA(A/G)(A/G)

TAF145p2F 5' GA(C/T)CC(A/T)AC(T/C/A)GG(A/T)TGTGG(A/T)GAAGG

TAF145p3R CT(G/A)GG(T/A)TGACC(T/A)ACACC(T/A)CTTCC 5'

TAF145p4R CC(A/T)GT(A/G)TAITC(T/C)AC(T/C/A)TT(G/A)TT 5'

FIG. 2B

Forward
Reverse
p1F p1F p2F
p3R p4R p4R
2000
1500
1000
750
500
250
• Template:C.albicans genomic DNA
• Major bands cloned and sequenced
* C. albicans TAF145 gene fragment

SUBSTITUTE SHEET (RULE 26)

FIG. 3A

```
   1 aaaaaaaaaataagctcccccgaaaaactgaaaccaaacaaacaaaaaagtaaaacaaacctcctacaaaactcttatat   80

  81 tatatccttgttacatttatatacgccaacaaccaaggataatacatttatgtatgataatgccagaagtcacacataaa  160

 161 tcaaacaatatagaagaagatgaagacaaagcatacaatcaatatgtaaatagtactgatattccagatgatcaaattat  240

 241 agaatcatatttcagaacgaaaaactgcccatgctgaagatgccatcgattatgaagatatagacgaattggcagaagaa  320

 321 gaagatgtg   ATG GAG GAT CTA CCC AGG GAT GAG GCA ATA AAT GGA TTA AAC TCA AAC AAC   380
   1              M   E   D   L   P   R   D   E   A   I   N   G   L   N   S   N   N    17

 381 AAC AAC AAC CAC GAT AAA GAT GAT GAT GAC GAC GAT GAA TTC AAT AGG TTA CTT CAA GAA   440
  18  N   N   N   H   D   K   D   D   D   D   D   D   E   F   N   R   L   L   Q   E    37

 441 GGG CAG CCT GAA TTG ACA AAT GAT GAA GAA ATG GCA GCT CAA GCT GCT GCT GAA TCT CAA   500
  38  G   Q   P   E   L   T   N   D   E   E   M   A   A   Q   A   A   A   E   S   Q    57

 501 TTT GAT GCT TTG TTT GGA AAC TCT AAT GAT TTT GAT AGT AAT ATT AGC CAC CAT GAT CAT   560
  58  F   D   A   L   F   G   N   S   N   D   F   D   S   N   I   S   H   H   D   H    77

 561 ATG GGC GGG GAC AGT AAT GGT ATT ATT GAT GAT AAT CAC CAC AGC AGT GTA AAC GAC CAC   620
  78  M   G   G   D   S   N   G   I   I   D   D   N   H   H   S   S   V   N   D   H    97

 621 GAT GGT TTA TTC AAC AAT TTA GGA AAT GGT AAT CAT TTG CTA GAT GAT GAT AAT GAT GGC   680
  98  D   G   L   F   N   N   L   G   N   G   N   H   L   L   D   D   D   N   D   G   117

 681 TTG AAT GAT TTA GGT GAA CTA TTT GAT GAT CAA CAA GAG GAC AGC AAT GTT ATC AAC ACT   740
 118  L   N   D   L   G   E   L   F   D   D   Q   Q   E   D   S   N   V   I   N   T   137
```

A ——————————————————————————————— A

09/601965 PCT/US99/02940

FIG. 3B

```
A
741 AAG AAG CAT AAG CTA GAT GAC GAT AGC AAC AAC GAT GGC AAG ACT GCT CAA GAA GAT CAA  800
138 K   K   H   K   L   D   D   D   S   N   N   D   G   K   T   A   Q   E   D   Q    157

801 AAA GAG AAA GAA AAT AAA CGA CAA TTG AAA CGG CAA AAA CTA CAA AAG ATT GTT AAA CAT  860
158 K   E   K   E   N   K   R   Q   L   K   R   Q   K   L   Q   K   I   V   K   H    177

861 CTT GAG AAG GAA CAA ATC AAA CGA AAT ATA AAA TAT TAT TTC CCT ACT TAT TCA AGA CAT  920
178 L   E   K   E   Q   I   K   R   N   I   K   Y   Y   F   P   T   Y   S   R   H    197

921 AGA CCA TTT AAT TTC CAC AAA TTT TTT TCA CCA AGT CCT CAG TAT TAC CGT TAT CAA AGA  980
198 R   P   F   N   F   H   K   F   F   S   P   S   P   Q   Y   Y   R   Y   Q   R    217

981 CCA GCA ATT GCC TTG TCG AAA AAT ATA AAA CCA TTA ATA CCT ACA AAA GTA AAT CTT GAA 1040
218 P   A   I   A   L   S   K   N   I   K   P   L   I   P   T   K   V   N   L   E    237

1041 ATC GAG GTT GAT CAA AAG AAA ATT TTC AAA TTA AGA AGT GCT GAT ACT GCA TCG TTG TCA 1100
238  I   E   V   D   Q   K   K   I   F   K   L   R   S   A   D   T   A   S   L   S    257

1101 CAC GAA GAC AAA AAT GTC ACC AAT ATT ACT CAA GAT GAC TTG GAT TTT ATC AAA AAT TTA 1160
258  H   E   D   K   N   V   T   N   I   T   Q   D   D   L   D   F   I   K   N   L    277

1161 GAA AGC AAA AGA TCT TCT ATT GAC TCG TTT ATT AAA GAA ATT GAT TAC GTT AAA CGT GAT 1220
278  E   S   K   R   S   S   I   D   S   F   I   K   E   I   D   Y   V   K   R   D    297

1221 TGG ACT AAT TGC GAC AAG TTT GAT CAT TAT TCG AAA GAT TTA GTT CTA TCT ACC ACT GAT 1280
298  W   T   N   C   D   K   F   D   H   Y   S   K   D   L   V   L   S   T   T   D    317
B
```

09601965 .020700

FIG. 3C

```
B                                                                                    B
1281 TGG GAT GAT GAT GCT ATT ATA AAT GCT GGA GAC AAT GAG TAC TCT ATT GTG AAG CCA ATC 1340
 318  W   D   D   D   A   I   I   N   A   G   D   N   E   Y   S   I   V   K   P   I   337

1341 AAT GAG CTT TTG CTC AAC AAT CCC TTG GAC AAT AGT AAA CAG AAT AGA CAA AAA ATC GAG 1400
 338  N   E   L   L   L   N   N   P   L   D   N   S   K   Q   N   R   Q   K   I   E   357

1401 AAT GAC AAT ACT ACC AAC AAC TAT AAC CAA AAC AAT AGT AAT GTC CAA GAT GAG GAG GAG 1460
 358  N   D   N   T   T   N   N   Y   N   Q   N   N   S   N   V   Q   D   E   E   E   377

1461 GAT GAT GAT ATC TTC AAT GGA CAA ATA AAC TTG GAT AAA TTG AAA CTT GAT ATG AAT GAT 1520
 378  D   D   D   I   F   N   G   Q   I   N   L   D   K   L   K   L   D   M   N   D   397

1521 CCT AAC TTG TTA TTT GTT CCT AGT AAA AAA GTC GAT GCT ACC AAA TCA GTG GTT CCA AGT 1580
 398  P   N   L   L   F   V   P   S   K   K   V   D   A   T   K   S   V   V   P   S   417

1581 ACA GAT AAA TTA TTA GAA TTA AAG TTC AAC ATA TCT AAC GAT CAA GAG TAT GAA TTA TTG 1640
 418  T   D   K   L   L   E   L   K   F   N   I   S   N   D   Q   E   Y   E   L   L   437

1641 AGA AAG AAT TAC AAC ACC AAA CAA AGA TCT CAA TTG AGT AAT CTT AAT ATT GAA CAT TCA 1700
 438  R   K   N   Y   N   T   K   Q   R   S   Q   L   S   N   L   N   I   E   H   S   457

1701 GTT CCC GCA TTG CGA TTA CAG ACA CCT TAT TAT AAA GTC AAA CTT AGC ACA GAT GAA ACG 1760
 458  V   P   A   L   R   L   Q   T   P   Y   Y   K   V   K   L   S   T   D   E   T   477

1761 AGA TCA TTC CAT CGA CCA GTG TTT AAT GTC AGA CCT GGT ACA TTG GTG AGC TTT TCT AAA 1820
 478  R   S   F   H   R   P   V   F   N   V   R   P   G   T   L   V   S   F   S   K   497

1821 TTG AAG TTG CGG AAG CGG AAA AAA GAC AAG GGG AAA TCT TTG CAA CAG ATT TTT TCC AAA 1880
 498  L   K   L   R   K   R   K   K   D   K   G   K   S   L   Q   Q   I   F   S   K   517
C                                                                                    C
```

FIG. 3D

C —— C

1881	ACT	AGT	GAC	TTG	ACA	GTT	GCT	GAT	ACT	GGT	AAT	ATA	ATT	GCC	TTA	GAG	TAT	TCC	GAA	CAG	1940
518	T	S	D	L	T	V	A	D	T	G	N	I	I	A	L	E	Y	S	E	Q	537
1941	TAT	CCA	CCA	ATT	TTA	TCG	AAT	TTC	GGG	ATG	GGG	TCG	AAA	TTG	ATC	AAC	TAT	TAT	CGT	AAA	2000
538	Y	P	P	I	L	S	N	F	G	M	G	S	K	L	I	N	Y	Y	R	K	557
2001	GAA	AGA	CCA	AAC	GAC	ACT	TCA	CGT	CCC	AAG	GCT	CAA	ATT	GGG	GAA	ACT	CAT	ATT	TTG	GGG	2060
558	E	R	P	N	D	T	S	R	P	K	A	Q	I	G	E	T	H	I	L	G	577
2061	GTG	GAG	GAT	AGA	TCC	CCA	TTT	TGG	AAT	TTC	GGT	GAA	GTT	GCT	CCT	GGA	GAT	TTT	GTT	CCC	2120
578	V	E	D	R	S	P	F	W	N	F	G	E	V	A	P	G	D	F	V	P	597
2121	ACA	TTG	TAT	AAT	AAT	ATG	GTA	AGA	GCA	CCA	ATT	TTC	AAG	CAT	GAC	AAC	AAA	CCA	ACT	GAT	2180
598	T	L	Y	N	N	M	V	R	A	P	I	F	K	H	D	N	K	P	T	D	617
2181	TTT	CTA	TTG	GTT	AAA	TCT	CAA	GGG	GCT	GGA	TCG	CAC	CAA	AAA	TTT	TAT	TTG	CGA	GGG	ATC	2240
618	F	L	L	V	K	S	Q	G	A	G	S	H	Q	K	F	Y	L	R	G	I	637
2241	AAT	TTC	AAC	TTT	GCT	GTT	GGT	AAC	ACA	TTC	CCA	GTT	GAA	GTT	CCA	GCT	CCT	CAC	TCG	AGA	2300
638	N	F	N	F	A	V	G	N	T	F	P	V	E	V	P	A	P	H	S	R	657
2301	AAA	GTG	ACA	AAT	ATC	TCA	AAA	AAT	AGG	TTG	AAA	ATG	GTA	GTT	TTC	AGA	GTG	ATG	AAT	AGC	2360
658	K	V	T	N	I	S	K	N	R	L	K	M	V	V	F	R	V	M	N	S	677

D —— D

FIG. 3E

D ——— D

```
2361 TTA GGA GTG CCA CGT ATT TCT GTG AAA GAT GTT TCC AAA CAT TTT CCT GAG CAT AGC GAT 2420
 678  L   G   V   P   R   I   S   V   K   D   V   S   K   H   F   P   E   H   S   D   697

2421 ATG CAA AAT CGA CAA AGA TTG AAA GAA TTT ATG GAA TAC CAA AGA CAA GGT GAG GAC CAA 2480
 698  M   Q   N   R   Q   R   L   K   E   F   M   E   Y   Q   R   Q   G   E   D   Q   717

2481 GGG TAT TGG AAA GTA AGA GGA TTA AAC GAT GTA ATT CCA GGA GAA GAG GAA ATT CGA ACT 2540
 718  G   Y   W   K   V   R   G   L   N   D   V   I   P   G   E   E   E   I   R   T   737

2541 ATG ATT ACA CCA GAA GAT TCT TCA TTG ATG GAC ACA ATG CAA TTT GGA CAG CAA GTA CTA 2600
 738  M   I   T   P   E   D   S   S   L   M   D   T   M   Q   F   G   Q   Q   V   L   757

2601 GAT GAT AAT ATG GTG TTA TTC GGT GAG CAG AGC AGA CAA GAA TCA TCT AGA AGT AGA AAA 2660
 758  D   D   N   M   V   L   F   G   E   Q   S   R   Q   E   S   S   R   S   R   K   777

2661 GGT GAT AAA AGA GAG GAT TCC ATT GCG GAT GAT GCA GAA AAT GGA GAT GAT ATA AAT AAG 2720
 778  G   D   K   R   E   D   S   I   A   D   D   A   E   N   G   D   D   I   N   K   797

2721 GAT AAA GAG AAA GAG GTT GAG AAA GAA AAG GAA CAG GAA AGA GAA GAA GAG AAA GGG AAA 2780
 798  D   K   E   K   E   V   E   K   E   K   E   Q   E   R   E   E   E   K   G   K   817

2781 GAT AAG GAA AAA GAC AAG GAC AAG GAA AAA GAC AAA ACC GAG AAG GAG AAA TCA AAG AAA 2840
 818  D   K   E   K   D   K   D   K   E   K   D   K   T   E   K   E   K   S   K   K   837

2841 TCA AAG GAG CAA GAC ACT GAA ATT GAT GTT GAA GAA GAA TTG GCA CCA TGG AAC TTA TCG 2900
 838  S   K   E   Q   D   T   E   I   D   V   E   E   E   L   A   P   W   N   L   S   857

2901 AGA AAT TTT GTT ATC GCC AAT CAG ACC AAG ACA ATG CTA CAA TTG AAT GGT GAA GGT GAT 2960
 858  R   N   F   V   I   A   N   Q   T   K   T   M   L   Q   L   N   G   E   G   D   877
```

E ——— E

09601965 .102000

FIG. 3F

E

```
2961 CCA ACT GGA ATT GGG TTA GGA TTT TCC ATG TTG AGG GCT ACA CAA AAG AAC CCA TTC AAA 3020
 878  P   T   G   I   G   L   G   F   S   M   L   R   A   T   Q   K   N   P   F   K  897

3021 CCG TTA TTT ACC CCA CCA CCA GAA AAT GTC CCT AAA AGT AAT GCT GCA GCC CAT AAT CAA 3080
 898  P   L   F   T   P   P   P   E   N   V   P   K   S   N   A   A   A   H   N   Q  917

3081 AAG TTG TAC GAA CAA GAG ATA AAA AGA ATA TGG TAC TCT CAA AGA AGC TCT TTA GTT GAT 3140
 918  K   L   Y   E   Q   E   I   K   R   I   W   Y   S   Q   R   S   S   L   V   D  937

3141 CAT GGG GAA GGA ACT GAA TCA AAG TTG CAA CAG ATC TAT AAT GAG TAC CCG CCA GCA GAT 3200
 938  H   G   E   G   T   E   S   K   L   Q   Q   I   Y   N   E   Y   P   P   A   D  957

3201 CAT GAA TTG TAT TTG AAA AAC AAA CTT GAA CAA GAC CAA CAG GTA CAA CAA CAA CAA CAA 3260
 958  H   E   L   Y   L   K   N   K   L   E   Q   D   Q   Q   V   Q   Q   Q   Q   Q  977

3261 GAT CCT CTG CTT CAG GCT GAT CAG CAA CAG CAA CAA CAA CAA CAA CAG AAT CGA GTT TTG 3320
 978  D   P   S   L   Q   A   D   Q   Q   Q   Q   Q   Q   Q   Q   Q   N   R   V   L  997

3321 AGG ATT ACT AGA AGA GTG CGA GAT GAA AAT GGA ATA GTG CAT AGA AAA GTT GAA TTT ATT 3380
 998  R   I   T   R   R   V   R   D   E   N   G   I   V   H   R   K   V   E   F   I 1017

3381 CAT GAT CCA AGA TTA ATT AGA GCA TAT GTT AAG CGT AAG AAA CAA ATT GAA GAT GAG TTA 3440
1018  H   D   P   R   L   I   R   A   Y   V   K   R   K   K   Q   I   E   D   E   L 1037

3441 TTG AAG AAT GCT GAT GTT GAT GAA ATA TTA CCT ACT AAC GAC AAG GAG TTA AAC AAA ATC 3500
1038  L   K   N   A   D   V   D   E   I   L   P   T   N   D   K   E   L   N   K   I 1057
```

F

SUBSTITUTE SHEET (RULE 26)

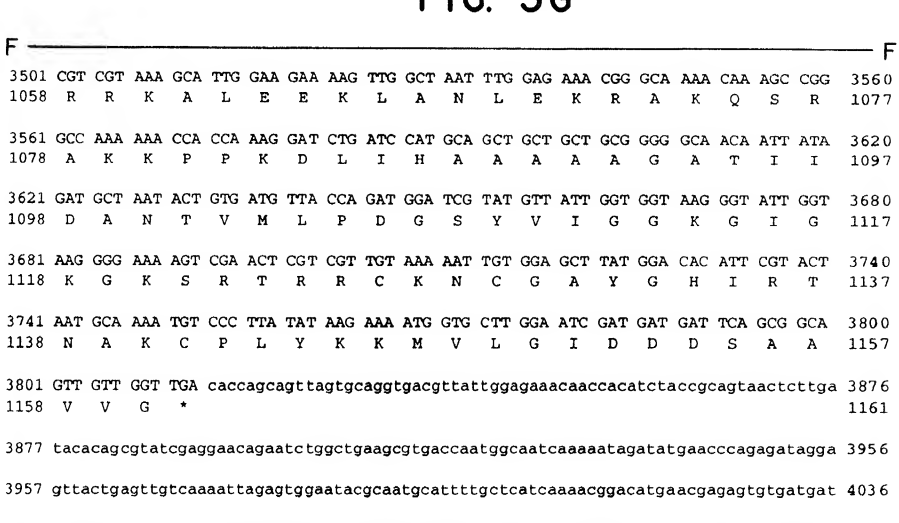
WO 99/40199

09/601965

PCT/US99/02940

000201 S9610960

FIG. 3G



SUBSTITUTE SHEET (RULE 26)

WO 96/01961

09/601965

PCT/US99/02940

FIG. 4A

```
C. albicans    MEDLPRDEAINGLNSNNNNNNHDKDDDDDDEFNRLLQEGQPELTNDEEMAAQAAAESQFDALFGNSNDFDS 70
                .|   :.   . :  :. : :.:|: .|: :  | :|::::|  .     :. .. :: :|::| .:: .
S.cerevisiae   GGDEGANSKDYTEHLPDAVDFEDEDELADDDDDLPEESDANLHPAMMIMGAYDDVNENGAVLGIDSN-SL 105

C. albicans    NISHHDHMGGDSNGIIDDNHHSSVNDHDGLFNNLGNGNHLLDDDNDGLNDLGELFDDQQEDSNVINTKKH 140
               |:  .:   |: |: :| ::: :: .. ::|| .::..:  |..:.: |:: |    .:: : : ::.
S.cerevisiae   NMQLPEINGDLSQQFILEDDGGTPATSNALFMGMDANEIHLATEIGVLDGSG----ANEIGHSQLSIGGV 171

C. albicans    KLDDDSNNDGKTAQEDQKEKENKRQLKRQKLQKIVKHLEKEQIKRNIKYYFPTYSRHRPFNFHKFFSPSP 210
               : :| | |:| . : | .: .:|:.     ||: :::| ||  :|   || ||.:.: : ::.:|:: ..:
S.cerevisiae   NGNDMSINGGFIMEPDMSDGKHKKA---TKLD-LINH-EKYLLK---KY-FPDFEKGKILKWNKLIYRRS 232

C. albicans    QYYRYQRPAIALSKNIKPLIPTKVNLEIEVDQKKIFKLRSADTASLSHEDKNVINITQDDLDFIKNLESK 280
                 .|::.: .:|:  ||::| :::::.:: |:|::|: |: . .:  .::||  |: |.: .  :. .
S.cerevisiae   --VPYHWHS-EISRVKKPFMPLNLKFKVQQDDKRLFNSRTISYVAPIYQGKN--NLLQSNSSASRR-GLI 296

C. albicans    RSSIDSFIKEIDYVKRDWINCDKFDHYSKDLVLSTTDWDDDAIINAGDNEYSIVKPINELLLNNPLDNSK 350
               : |||.::  |.  ::.:.  :. .  |.||:::|.|||::  |||.|.:. :.:   :     :  :. |
S.cerevisiae   HVSIDELFP-IKEQQKKRKIIHDEKTISEDLLIATDDWDQEKIINQGTSSTATLADSSMTPNLKFSGGYK 365
```

A —————————————————————————— A

FIG. 4B

A ——— A

```
C. albicans    QNRQKIENDNTINN 364
                :.   : .:. :
S.cerevisiae   LKSLIEDVAEDWQW 379

C. albicans    EDDDIFNGQINLDK-LKLDMNDPNLLFVPSK---------KVDATKSVVPSTDKLLELKFNISNDQEYEL 436
               ::| |:::::: .|  .|:||| :||:: .|         ::|::: ::| .:.:|: |||:||||:.|::
S.cerevisiae   DEDMIIDAKLKESKHAELNMNDEKLLLMIEKTNNLAQQKQQLDSSNLILPLNETTLQQKFNLSNDDKYQI 449

C. albicans    LRKNYNTKQRSQLSNLNIEHSVPALRLQTPYYKVKLSTDETRSFHRPVF--NVRPGTLVSFSKLKLRKRK 504
               |:|..:|| || :||||||:|| ||:.||:|:||| ::   : | |||  | ::|||| : ||||| ||||
S.cerevisiae   LKKTHQTKVRSTISNLNIQHSQPAINLQSPFYKVAVPRYQLRHFHRENFGSHLRPGTKIVFSKLKARKRK 519

C. albicans    KDKGKSLQQIFSKTSDLTVADTGNIIALEYSEQYPPILSNFGMGSKLINYYRKERPNDTSRPKAQIGETH 574
               :||||.::: ||.: |||::||: :  :||||| |  ||:|||::||||||||.. :|| ||| .:||||
S.cerevisiae   RDKGKDVKESFSTSQDLTIGDTAPVYLMEYSEQTPVALSKFGMANKLINYYRKANEQDTLRPKLPVGETH 589

C. albicans    ILGVEDRSPFWNFGEVAPGDFVPTLYNNMVRAPIFKHDNKPTDFLLVKSQGAGSHQKFYLRGINFNFAVG 644
               :|||:|:||||||| |.||::|||||||||:|||:|||| . |||||.|| | |  ::||||.|| |:||
S.cerevisiae   VLGVQDKSPFWNFGFVEPGHIVPTLYNNMIRAPVFKHDISGTDFLLTKSSGFGISNRFYLRNINHLFTVG 659
```

B ——— B

09601965 .100200

FIG. 4C

```
B ------------------------------------------------------------------------------ B

C. albicans    NTFPV-EVPAPHSRKVINISKNRLKMVVFRVMNSLGVPRISVKDVSKHFPEHSDMQNRQRLKEFMEYQRQ 713
               :||||  |:|:|:||||||::.  .||||:::|::|       ||:.  ::|||||::.   ||||::||||.|||:
S.cerevisiae   QTFPVEEIPGPNSRKVTSMKATRLKMIIYRILNHNHSKAISIDPIAKHFPDQDYGQNRQKVKEFMKYQRD 729

C. albicans    GEDQGYWKVRGLNDVIPGEEEIRIMITPEDSSLMDIMQFGQQVLDDNMVLFGEQSRQE 771
               | ::|  |:::: :: : ::|.::::||||: | :::|  | |  :||  .   : . :.
S.cerevisiae   GPEKGLWRLKD-DEKLLDNEAVKSLITPEQISQVESMSQGLQFQEDNEAYNFDSKLKS 786

C. albicans    VEEELAPWNLSRNFVIANQTKTMLQLNGEGDPTGIGLGFSMLRATQKNPFKPLFTPPPENVPKSNAAAHN 916
               :||:| |||:::||: :.| ::|:|::| ||||| | |||.|::: |.  |          | .::::::|
S.cerevisiae   LEENLLPWNITKNFINSTQMRAMIQIHGVGDPTGQGEGFSFLKTSMKGGF----------VKSGSPSSNN 846

C. albicans    QKLYEQEIKRIWYSQRSSLVDHGEGTESKLQQIYNEYPPADHELYLKNKLEQDQQVQQQQQDPSLQADQQ 986
               :.   .:..:   |       ::.:: .: :..:.:..   |..:. .| :.| :|      :  :.|   |::::
S.cerevisiae   NSSNKKGINIHSY----NVAQQQKAYDEEIAKIW--YIHIK-SLSISNPFE-----EMINPDEINQINKH 904

C ------------------------------------------------------------------------------ C
```


09601965 .10200

09/601965

09601965.101200

16/32

FIG. 4D

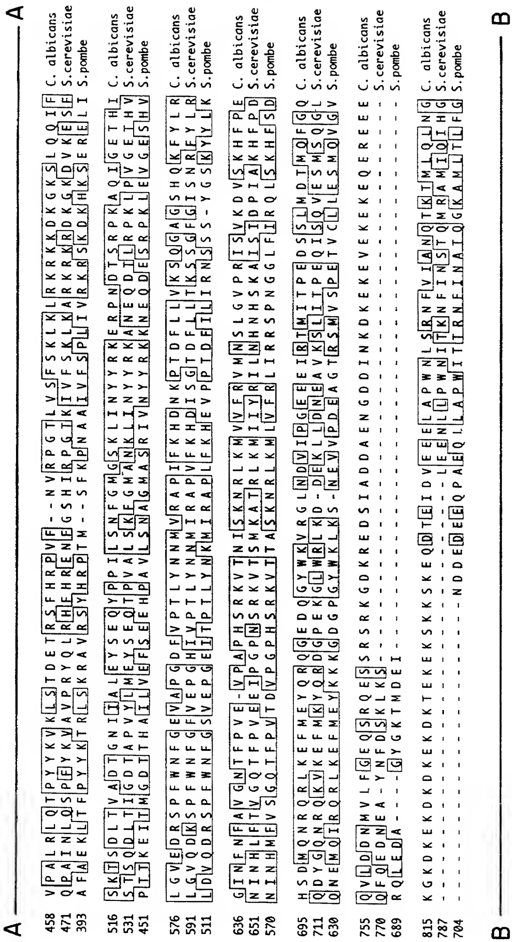
C —————————————————————————— C

```
C. albicans    QQQQQQQNRVLRITRRVRDENGIVHRKVEFIHDPRLIRAYVKRKKQIEDELLKNADVDEILPTNDKELNK 1056
                : ::::::|:|.|:  |||||||::|:. ||:|||:|::|:|  |.| .:::  |   : |.   .| .||:|
S.cerevisiae   VKTDRDDKKILKIVRKKRDENGIIQRQTIFIRDPRVIQGYIKIKEQDKEDVNKLLEEDTSKINNLEELEK 974

C. albicans    IRRKALEEKLANLEKRAKQSRAKKPPKDLIHAAAAAGATIIDANIVMLPDGSYVIGGKGIGKGKSRTRRC 1126
                ::| |: .||||||| ::|.|| :. ..  :::|: ... :|..:  |:: :   :::.   |||: ||||
S.cerevisiae   -QKKLLQLELANLEK-SQQRRAARQNSK-RNGGATRTENSVDNGSD-LAGVTDGKAARN--KGKNTTRRC 1038

C. albicans    KNCGAYGHIRTNAKCPLYKKMVLGIDDD 1154
                .||. |||||| .||:|...    . .
S.cerevisiae   ATCGQIGHIRTNKSCPMYSSKDNPASPK 1066
```

FIG. 5A

09/601965

FIG. 5B



09601965 . 020700

FIG. 5C

B —— B

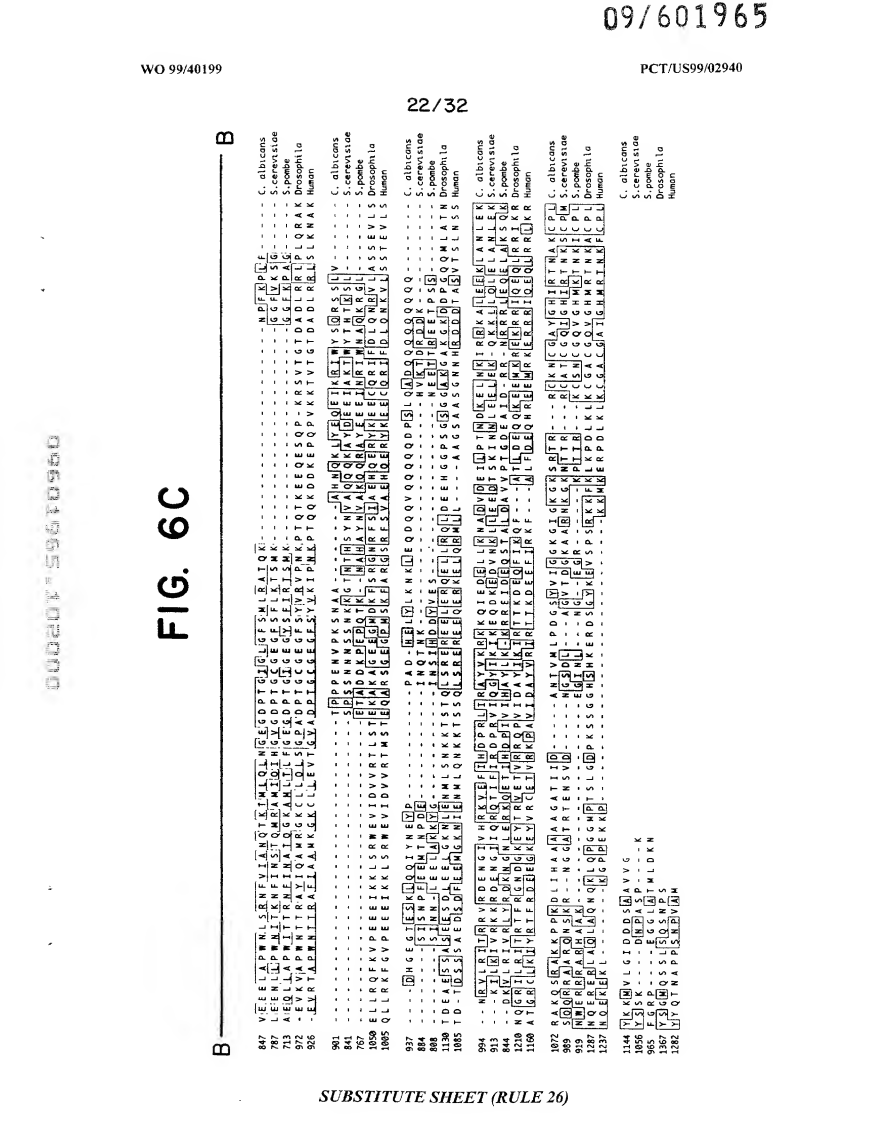
```
875  EGDPTGIGLGFSMLRATQKNPFKPLFTPPPENVPKSN---------AAAHNQKLYEQEIKR C. albicans
815  VGDPTGCGEGFSFLKTSMKGGFVKSGSPSSNNNSSNKKGTNTHSYNVAQQQKAYDEEIAK S.cerevisiae
741  EGDPTGIGEGYSFIRTSMKGGFKPAGETADDKPEPQTK--NAHAYNVAKQQRAYEEEINR S.pombe

927  IWYSQRSSLVDHGEGTESKLQQIYNEYPPADHELYLKNKLEQD---QQVQQQQQDPSLQA C. albicans
875  TWYTHTKSL----------SISNPFEE---------------MTNP---DEINQ--------T S.cerevisiae
799  IWNAQKRGL----------SINNL-EE---------------LAKKYGINSIHD--------D S.pombe

984  --DQQQQQQQQQ----NRVLRITRRVRDENGIVHRKVEFIHDPRLIRAYVKRKKQIEDEL C. albicans
902  --NKHVKTDRDD----KKILKIVRKKRDENGIIQRQTIFIRDPRVIQGYIKIKEQDKEDV S.cerevisiae
828  YVESNEETTREETPSSDKVLRIVRLYRDKNGNLERKQETIHDPIVIHAYLKKRREIDEQ- S.pombe

1038 LKNADVDEILPTNDKELNKIRRKALEEKLANLEKRAKQSRAKKPPKDLIHAAAAAGATII C. albicans
956  NKLLEEDTSKINNLEELEK-QKKLLQLELANLEKSQQRRAARQNSKR--NGGATRTENSV S.cerevisiae
887  --STALDAVVPTGDEAIDRRNRRRLEQELAKSQKNWERRRARHAAKE------------- S.pombe

1098 DANTVMLPDGSYVIGGKGIGKGKSRTRRCKNCGAYGHIRTNAKCPLYKKMVLGIDDDSAA C. albicans
1013 DNGSDL---AGVTDGKAARNKGKNTTRRCATCGQIGHIRTNKSCPMYSS------KDNPA S.cerevisiae
932  -----------GINLNGEGR---KPTTRKCSNCGQVGHMKTNKICPLFGR------PEGGL S.pombe

1158 VV----G                                                      C. albicans
1064 SP----K                                                      S.cerevisiae
973  ATMLDKN                                                      S.pombe
```


09601965 .102000

FIG. 6A

FIG. 6B

FIG. 6C



30° C

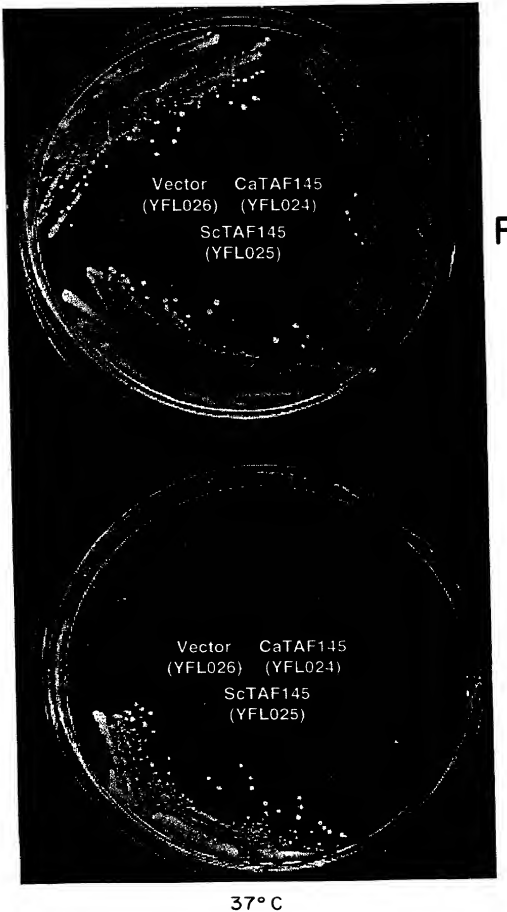


FIG. 7A

37° C

09601965 020299

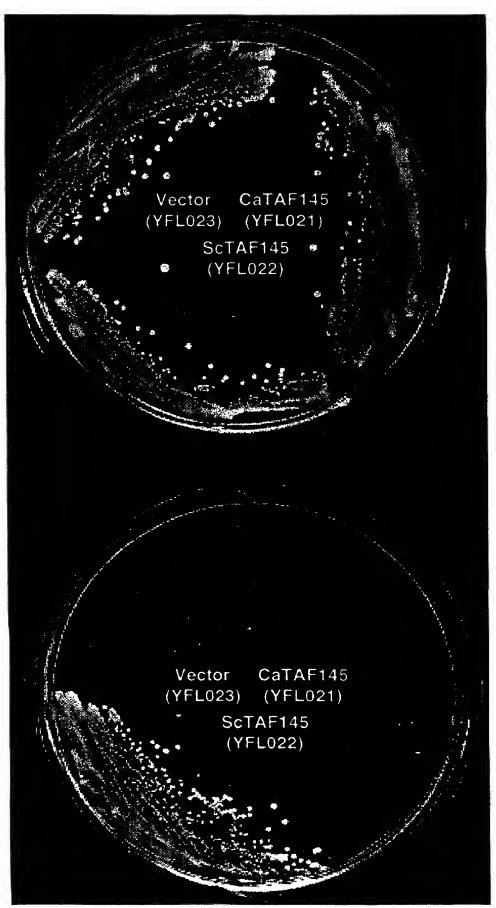

SC – Leu
Vector (YFL023)
CaTAF145 (YFL021)
ScTAF145 (YFL022)
Vector (YFL023)
CaTAF145 (YFL021)
ScTAF145 (YFL022)
SC – Leu + 5 – FOA

FIG. 7B

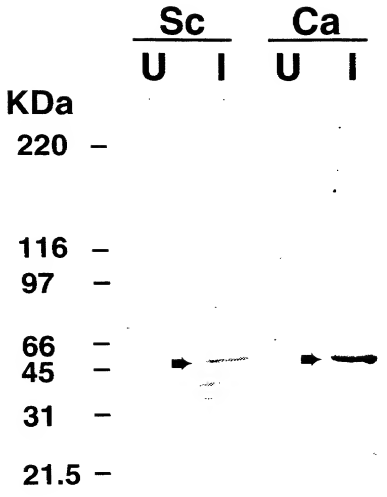
FIG. 8A

339 C. albicans
354 S.cerevisiae
288 S.pombe
432 Human
369 C. albicans
381 S.cerevisiae
305 S.pombe
482 Human
412 C. albicans
425 S.cerevisiae
346 S.pombe
532 Human
445 C. albicans
458 S.cerevisiae
380 S.pombe
580 Human
488 C. albicans
503 S.cerevisiae
423 S.pombe
630 Human
531 C. albicans
546 S.cerevisiae
466 S.pombe
680 Human
A
A

FIG. 8B

A
581 R S P F W N F G E V A P G D E V P T L Y N N M V R A P T F K H D N K P T D F L L V K S Q G A G S H Q C. albicans
596 K S P F W N F G E V E P G H I V P T L Y N N M I R A P V F K H D I S G T D F L L I K S S G F G I S N S.cerevisiae
516 R S P F W N F G S V E P G E I T P I L Y N K M I R A P L F K H E V P P T D F I L T R N S S - S Y G S S.pombe
727 T S P F - - L G S L H P G Q L L Q A F E N N L F R A P I Y L H K M P E T D F L I I R T R - - - - Q Human
631 K F Y L R G I N F N F A V G N T F P V E L - V P A P H S R K V T N I S K N R L K M V V E R V M - - N S C. albicans
646 R E Y L R N I N H L F T V G Q T F P V E E I P G P N S R K V T S M K A T R L K M I I Y R I L - - N H S.cerevisiae
565 K Y Y L K N I N H M F V S G Q T F P V T D V P G P H S R K V T T A S K N R L K M L V F R L I - - R R S.pombe
770 G Y Y I R E L V D I E V V G Q Q C P L F E E V P G P N S K R A N T H I R D F L Q V F I Y R L F W K S K Human
678 L G V P R I S V K D V S K H F P E H S D M Q N R Q R L K E F M E Y Q R Q G E D Q G Y W K V R G L N D C. albicans
694 N H S K A I S I D P I A K H F P D Q D Y G Q N R Q K V K E F M K Y Q R D G P E K G L W R L K D - D E S.cerevisiae
613 S P N G G L F I R Q L S K H F S D Q N E M Q I R Q R L K E F M E Y K K K G D G P G Y W K L K S - N E S.pombe
820 D R P R R I R M E D I K K A I E P S H S E S S I R K R L K L C A D F K R T G M D S N W V L K S - D F F Human
728 V I P G E E E I R T M I T P E D S S L M D I M Q F G Q Q V L D D N M V L - F G E Q C. albicans
743 K L L D N E A V K S L I I P E Q I S Q V E S M S Q G L L Q F Q E D N E A Y N E D S K - - L K S L S.cerevisiae
662 V V P D E A G I R S M V S P E T V C L L E S M Q V G V R Q L E D A - - - G Y G K T - - M D E S.pombe
869 R L P T E E E I R A M V S P E Q C C A Y Y S M I A A E Q R L K D A - - - G Y G E K S F F A P E Human
C.albicans
S.pombe
S.cerevisiae
Human
50
40
30
20
10
0

FIG. 9



09601965.102200

FIG. 10A

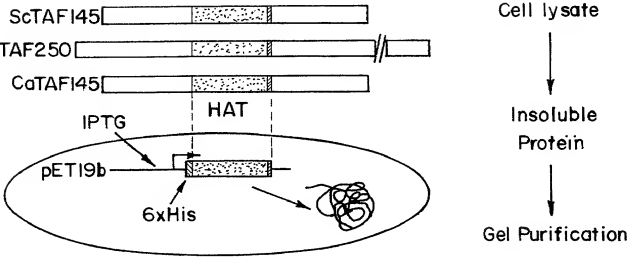


FIG. 10B

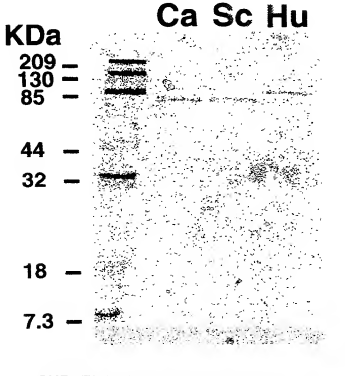


FIG. 11

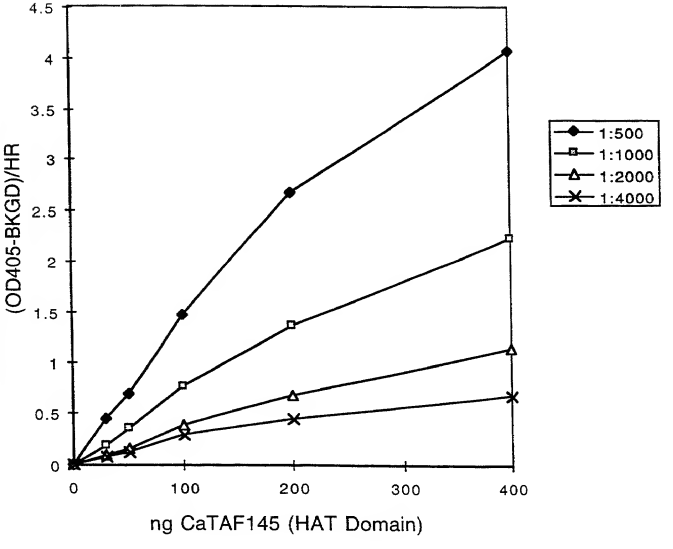

4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
(OD405-BKGD)/HR
0
100
200
300
400
ng CaTAF145 (HAT Domain)
1:500
1:1000
1:2000
1:4000

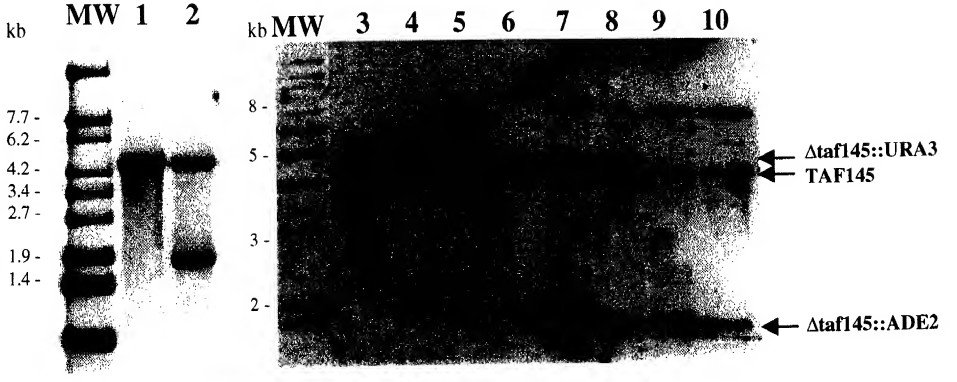
FIG. 12

MW
1
2
3
4
209
130
85
43
32
18
Coomassie
Western: Anti-C. albicans HAT Domain
1. Human TAF infected High-five cells
2. C. albicans TAF High-five infected cells
3. Mock-infected High-five cells
4. Non-infected High-five cells

FIG. 13A

CaADE2
TAF145 copy 1
TAF145 copy 2
CA18
(Ura-, Ade-)
Transformation #1
CaADE2
TAF145 copy 2
CaURA3
Δtaf145::ADE2 / TAF 145
(Ura-, Ade+)
Transformation #2
CaADE2
CaURA3
Δtaf145::ADE2 / TAF145::URA3 ???
(Ura+, Ade+)

09601965 .102000



Lane

1 **TAF145/TAF145 (CAI8)**

2 **Δtaf145::ADE2 / TAF145**

3 **TAF145 / Δtaf145::URA3**

4-10 **Δtaf145::ADE2 / TAF145, Ura+**

FIG. 13B

09/601965

SUBSTITUTE SHEET (RULE 26)

COMBINED DECLARATION FOR PATENT APPLICATION AND POWER OF ATTORNEY
(Includes Reference to PCT International Applications)

ATTORNEY DOCKET NUMBER
0342/1D516-US2

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below next to my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed for and which a patent is sought on the invention entitled:

A NOVEL FUNGAL PROTEIN CRITICAL FOR EXPRESSION OF FUNGAL PROTEINS

the specification of which (check only one item below):

OIPE
OCT 2 0 2000
PATENT & TRADEMARK OFFICE

[] is attached hereto.

[] was filed as United States application

Serial No. _____

on _____

and was amended

on _____ (if applicable).

[X] was filed as PCT international application

Number PCT/US99/02940

on 8 February 1999

and was amended under PCT Article 19

on _____ (if applicable).

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s) designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed:

PRIOR FOREIGN/PCT APPLICATION(S) AND ANY PRIORITY CLAIMS UNDER 35 U.S.C. 119:

COUNTRY (if PCT indicate PCT)	APPLICATION NUMBER	DATE OF FILING (day, month, year)	PRIORITY CLAIMED UNDER 35 U.S.C. 119
US	60/074,100	9 February 1998	[X] YES [] NO
			[] YES [] NO
			[] YES [] NO
			[] YES [] NO
			[] YES [] NO

US

Combined Declaration for Patent Application and Power of Attorney (Continued)
(Includes Reference to PCT International Applications)

ATTY'S DOCKET NUMBER
0342/1D516-US2

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) or PCT international application(s) designating the United States of America that is/are listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in that/those prior application(s) in the manner provided by the first paragraph of Title 35, United States Code §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application(s) and the national or PCT international filing date of this application:

PRIOR U.S. APPLICATIONS OR PCT INTERNATIONAL APPLICATIONS DESIGNATING THE U.S. FOR BENEFIT UNDER 35 U.S.C. 120:

U.S APPLICATIONS		STATUS (Check one)		
U.S APPLICATION NUMBER	U.S. FILING DATE	PATENTED	PENDING	ABANDONED

OIPE
OCT 2 0 2000
PATENT & TRADEMARK OFFICE

PCT APPLICATIONS DESIGNATING THE U.S.

PCT APPLICATION NO.	PCT FILING DATE	U.S SERIAL NUMBER ASSIGNED (if any)

POWER OF ATTORNEY: As a named inventor, I hereby appoint the following attorney(s) and/or agents to prosecute this application and transact all business in the Patent and Trademark Office connected therewith. Morris Relson #15,108, Gordon D. Coplein #19,165, William F. Dudine, Jr. #20,569, Michael J. Sweedler #19,937, S. Peter Ludwig #25,351, Paul Fields #20,298, Joseph B. Lerch #26,936, Melvin C. Garner #26,272, Ethan Horwitz #27,646, Beverly B. Goodwin #28,417, Adda C. Gogoris #29,714, Martin E. Goldstein #20,869, Bert J. Lewen 19,407, Henry Sternberg #22,408, Peter C. Schechter #31,662, Robert Schaffer #31,194, David R. Francescani #25,159, Robert C. Sullivan, Jr. #30,499, and Joseph R. Robinson #33,448

Send Correspondence to:

DARBY & DARBY P.C.
805 Third Avenue
New York, New York 10022-7513

Direct Telephone Calls to: (name and telephone number)

(212) 527-7700
Anne E. Zitron

2	FULL NAME OF INVENTOR	FAMILY NAME Thompson	FIRST GIVEN NAME Craig	SECOND GIVEN NAME M.
0	RESIDENCE & CITIZENSHIP	CITY Arlington	STATE OR FOREIGN COUNTRY Massachusetts MA	COUNTRY OF CITIZENSHIP U.S.A.
1	POST OFFICE ADDRESS	POST OFFICE ADDRESS 15 College Avenue	CITY Arlington	STATE & ZIP CODE/COUNTRY Massachusetts 02474
2	FULL NAME OF INVENTOR	FAMILY NAME Long	FIRST GIVEN NAME Fan	SECOND GIVEN NAME
0	RESIDENCE & CITIZENSHIP	CITY Waltham	STATE OR FOREIGN COUNTRY Massachusetts MA	COUNTRY OF CITIZENSHIP China, PRC
2	POST OFFICE ADDRESS	POST OFFICE ADDRESS 51 Garden Circle, Apt. #10 8148 Genesse Ave. Apt. 137	CITY Waltham San Diego	STATE & ZIP CODE/COUNTRY Massachusetts 02452 California 92122
2	FULL NAME OF INVENTOR	FAMILY NAME Wobbe	FIRST GIVEN NAME Richard	SECOND GIVEN NAME C.
0	RESIDENCE & CITIZENSHIP	CITY Lexington	STATE OR FOREIGN COUNTRY Massachusetts MA	COUNTRY OF CITIZENSHIP U.S.A.
3	POST OFFICE ADDRESS	POST OFFICE ADDRESS 57 Spring Street	CITY Lexington	STATE & ZIP CODE/COUNTRY Massachusetts 02173

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patents issuing thereon.

SIGNATURE OF INVENTOR 201	SIGNATURE OF INVENTOR 202	SIGNATURE OF INVENTOR 203
Craig Thompson	Tom Long	
DATE Aug. 24, 2000	DATE Oct 4, 2000	DATE Nov 25 2000

M:\WP\FORMS\WP\0396

SEQUENCE LISTING

<110> Scriptgen
Thompson, M. Craig
Long, Fan
Wobbe, C. Richard

<120> A NOVEL FUNGAL MULTISUBUNIT PROTEIN
COMPLEX CRITICAL FOR EXPRESSION OF FUNGAL PROTEINS

<130> 0342/2D516

<140> Unknown
<141> Filed Concurrently

<150> 60/074,100
<151> 1998-02-09

<160> 32

<170> FastSEQ for Windows Version 3.0

<210> 1
<211> 4116
<212> DNA
<213> C. albicans

<400> 1

aaaaaaaaaa taagctcccc cgaaaaactg aaaccaaaca aacaaaaaag taaaacaaac 60
ctcctacaaa actcttatat tatatccttg ttacatttat atacgccaac aaccaaggat 120
aatacattta tgtatgataa tgccagaagt cacacataaa tcaaacaata tagaagaaga 180
tgaagacaaa gcatacaatc aatatgtaaa tagtactgat attccagatg atcaaattat 240
agaatcatat ttcagaacga aaaactgccc atgctgaaga tgccatcgat tatgaagata 300
tagacgaatt ggcagaagaa gaagatgtga tggaggatct acccagggat gaggcaataa 360
atggattaaa ctcaaacaac aacaacaacc acgataaaga tgatgatgac gacgatgaat 420
tcaataggtt acttcaagaa gggcagcctg aattgacaaa tgatgaagaa atggcagctc 480
aagctgctgc tgaatctcaa tttgatgctt tgtttggaaa ctctaatgat tttgatagta 540
atattagcca ccatgatcat atgggcgggg acagtaatgg tattattgat gataatcacc 600
acagcagtgt aaacgaccac gatggtttat tcaacaattt aggaaatggt aatcatttgc 660
tagatgatga taatgatggc ttgaatgatt taggtgaact atttgatgat caacaagagg 720
acagcaatgt tatcaacact aagaagcata agctagatga cgatagcaac aacgatggca 780
agactgctca agaagatcaa aaagagaaag aaaataaacg acaattgaaa cggcaaaaac 840
tacaaaagat tgttaaacat cttgagaagg aacaaatcaa acgaaatata aaatattatt 900
tccctactta ttcaagacat agaccattta atttccacaa atttttttca ccaagtcctc 960
agtattaccg ttatcaaaga ccagcaattg ccttgtcgaa aaatataaaa ccattaatac 1020
ctacaaaagt aaatcttgaa atcgaggttg atcaaaagaa aattttcaaa ttaagaagtg 1080
ctgatactgc atcgttgtca cacgaagaca aaaatgtcac caatattact caagatgact 1140
tggattttat caaaaattta gaaagcaaaa gatcttctat tgactcgttt attaaagaaa 1200

ttgattacgt taaacgtgat tggactaatt gcgacaagtt tgatcattat tcgaaagatt 1260
tagttctatc taccactgat tgggatgatg atgctattat aaatgctgga gacaatgagt 1320
actctattgt gaagccaatc aatgagcttt tgctcaacaa tcccttggac aatagtaaac 1380
agaatagaca aaaaatcgag aatgacaata ctaccaacaa ctataaccaa aacaatagta 1440
atgtccaaga tgaggaggag gatgatgata tcttcaatgg acaaataaac ttggataaat 1500
tgaaacttga tatgaatgat cctaacttgt tatttgttcc tagtaaaaaa gtcgatgcta 1560
ccaaatcagt ggttccaagt acagataaat tattagaatt aaagttcaac atatctaacg 1620
atcaagagta tgaattattg agaaagaatt acaacaccaa acaaagatct caattgagta 1680
atcttaatat tgaacattca gttcccgcat tgcgattaca gacaccttat tataaagtca 1740
aacttagcac agatgaaacg agatcattcc atcgaccagt gtttaatgtc agacctggta 1800
cattggtgag cttttctaaa ttgaagttgc ggaagcggaa aaaagacaag gggaaatctt 1860
tgcaacagat tttttccaaa actagtgact tgacagttgc tgatactggt aatataattg 1920
ccttagagta ttccgaacag tatccaccaa ttttatcgaa tttcgggatg gggtcgaaat 1980
tgatcaacta ttatcgtaaa gaaagaccaa acgacacttc acgtcccaag gctcaaattg 2040
gggaaactca tattttgggg gtggaggata gatccccatt ttggaatttc ggtgaagttg 2100
ctcctggaga ttttgttccc acattgtata ataatatggt aagagcacca attttcaagc 2160
atgacaacaa accaactgat tttctattgg ttaaatctca aggggctgga tcgcaccaaa 2220
aattttattt gcgagggatc aatttcaact ttgctgttgg taacacattc ccagttgaag 2280
ttccagctcc tcactcgaga aaagtgacaa atatctcaaa aaataggttg aaaatggtag 2340
ttttcagagt gatgaatagc ttaggagtgc cacgtatttc tgtgaaagat gtttccaaac 2400
attttcctga gcatagcgat atgcaaaatc gacaaagatt gaaagaattt atggaatacc 2460
aaagacaagg tgaggaccaa gggtattgga aagtaagagg attaaacgat gtaattccag 2520
gagaagagga aattcgaact atgattacac cagaagattc ttcattgatg gacacaatgc 2580
aatttggaca gcaagtacta gatgataata tggtgttatt cggtgagcag agcagacaag 2640
aatcatctag aagtagaaaa ggtgataaaa gagaggattc cattgcggat gatgcagaaa 2700
atggagatga tataaataag gataaagaga aagaggttga gaaagaaaag gaacaggaaa 2760
gagaagaaga gaaagggaaa gataaggaaa aagacaagga caaggaaaaa gacaaaaccg 2820
agaaggagaa atcaaagaaa tcaaaggagc aagacactga aattgatgtt gaagaagaat 2880
tggcaccatg gaacttatcg agaaattttg ttatcgccaa tcagaccaag acaatgctac 2940
aattgaatgg tgaaggtgat ccaactggaa ttgggttagg attttccatg ttgagggcta 3000
cacaaaagaa cccattcaaa ccgttattta ccccaccacc agaaaatgtc cctaaaagta 3060
atgctgcagc ccataatcaa aagttgtacg aacaagagat aaaaagaata tggtactctc 3120
aaagaagctc tttagttgat catggggaag gaactgaatc aaagttgcaa cagatctata 3180
atgagtaccc gccagcagat catgaattgt atttgaaaaa caaacttgaa caagaccaac 3240
aggtacaaca acaacaacaa gatcctctgc ttcaggctga tcagcaacag caacaacaac 3300
aacaacagaa tcgagttttg aggattacta gaagagtgcg agatgaaaat ggaatagtgc 3360
atagaaaagt tgaatttatt catgatccaa gattaattag agcatatgtt aagcgtaaga 3420
aacaaattga agatgagtta ttgaagaatg ctgatgttga tgaaatatta cctactaacg 3480
acaaggagtt aaacaaaatc cgtcgtaaag cattggaaga aaagttggct aatttggaga 3540
aacgggcaaa acaaagccgg gccaaaaaac caccaaagga tctgatccat gcagctgctg 3600
ctgcgggggc aacaattata gatgctaata ctgtgatgtt accagatgga tcgtatgtta 3660
ttggtggtaa gggtattggt aaggggaaaa gtcgaactcg tcgttgtaaa aattgtggag 3720
cttatggaca cattcgtact aatgcaaaat gtcccttata taagaaaatg gtgcttggaa 3780
tcgatgatga ttcagcggca gttgttggtt gacaccagca gttagtgcag gtgacgttat 3840
tggagaaaca accacatcta ccgcagtaac tcttgataca cagcgtatcg aggaacagaa 3900
tctggctgaa gcgtgaccaa tggcaatcaa aaatagatat gaacccagag ataggagtta 3960
ctgagttgtc aaaattagag tggaatacgc aatgcatttt gctcatcaaa acggacatga 4020
acgagagtgt gatgatgtta ttgtcagtaa tataggttta gtttacattt taatgacata 4080
taaacaatgt aatgattatt atgcatttct atttgt 4116

<210> 2
<211> 1160
<212> PRT
<213> C. albicans

<400> 2

```
Met Glu Asp Leu Pro Arg Asp Glu Ala Ile Asn Gly Leu Asn Ser Asn
 1               5                  10                  15
Asn Asn Asn Asn His Asp Lys Asp Asp Asp Asp Asp Asp Glu Phe Asn
            20                  25                  30
Arg Leu Leu Gln Glu Gly Gln Pro Glu Leu Thr Asn Asp Glu Glu Met
        35                  40                  45
Ala Ala Gln Ala Ala Ala Glu Ser Gln Phe Asp Ala Leu Phe Gly Asn
    50                  55                  60
Ser Asn Asp Phe Asp Ser Asn Ile Ser His His Asp His Met Gly Gly
65                  70                  75                  80
Asp Ser Asn Gly Ile Ile Asp Asp Asn His His Ser Ser Val Asn Asp
                85                  90                  95
His Asp Gly Leu Phe Asn Asn Leu Gly Asn Gly Asn His Leu Leu Asp
            100                 105                 110
Asp Asp Asn Asp Gly Leu Asn Asp Leu Gly Glu Leu Phe Asp Asp Gln
        115                 120                 125
Gln Glu Asp Ser Asn Val Ile Asn Thr Lys Lys His Lys Leu Asp Asp
    130                 135                 140
Asp Ser Asn Asn Asp Gly Lys Thr Ala Gln Glu Asp Gln Lys Glu Lys
145                 150                 155                 160
Glu Asn Lys Arg Gln Leu Lys Arg Gln Lys Leu Gln Lys Ile Val Lys
                165                 170                 175
His Leu Glu Lys Glu Gln Ile Lys Arg Asn Ile Lys Tyr Tyr Phe Pro
            180                 185                 190
Thr Tyr Ser Arg His Arg Pro Phe Asn Phe His Lys Phe Phe Ser Pro
        195                 200                 205
Ser Pro Gln Tyr Tyr Arg Tyr Gln Arg Pro Ala Ile Ala Leu Ser Lys
    210                 215                 220
Asn Ile Lys Pro Leu Ile Pro Thr Lys Val Asn Leu Glu Ile Glu Val
225                 230                 235                 240
Asp Gln Lys Lys Ile Phe Lys Leu Arg Ser Ala Asp Thr Ala Ser Leu
                245                 250                 255
Ser His Glu Asp Lys Asn Val Thr Asn Ile Thr Gln Asp Asp Leu Asp
            260                 265                 270
Phe Ile Lys Asn Leu Glu Ser Lys Arg Ser Ser Ile Asp Ser Phe Ile
        275                 280                 285
Lys Glu Ile Asp Tyr Val Lys Arg Asp Trp Thr Asn Cys Asp Lys Phe
    290                 295                 300
Asp His Tyr Ser Lys Asp Leu Val Leu Ser Thr Thr Asp Trp Asp Asp
305                 310                 315                 320
Asp Ala Ile Ile Asn Ala Gly Asp Asn Glu Tyr Ser Ile Val Lys Pro
                325                 330                 335
Ile Asn Glu Leu Leu Leu Asn Asn Pro Leu Asp Asn Ser Lys Gln Asn
            340                 345                 350
```

```
Arg Gln Lys Ile Glu Asn Asp Asn Thr Thr Asn Asn Tyr Asn Gln Asn
        355                 360                 365
Asn Ser Asn Val Gln Asp Glu Glu Glu Asp Asp Asp Ile Phe Asn Gly
    370                 375                 380
Gln Ile Asn Leu Asp Lys Leu Lys Leu Asp Met Asn Asp Pro Asn Leu
385                 390                 395                 400
Leu Phe Val Pro Ser Lys Lys Val Asp Ala Thr Lys Ser Val Val Pro
                405                 410                 415
Ser Thr Asp Lys Leu Leu Glu Leu Lys Phe Asn Ile Ser Asn Asp Gln
            420                 425                 430
Glu Tyr Glu Leu Leu Arg Lys Asn Tyr Asn Thr Lys Gln Arg Ser Gln
        435                 440                 445
Leu Ser Asn Leu Asn Ile Glu His Ser Val Pro Ala Leu Arg Leu Gln
    450                 455                 460
Thr Pro Tyr Tyr Lys Val Lys Leu Ser Thr Asp Glu Thr Arg Ser Phe
465                 470                 475                 480
His Arg Pro Val Phe Asn Val Arg Pro Gly Thr Leu Val Ser Phe Ser
                485                 490                 495
Lys Leu Lys Leu Arg Lys Arg Lys Lys Asp Lys Gly Lys Ser Leu Gln
            500                 505                 510
Gln Ile Phe Ser Lys Thr Ser Asp Leu Thr Val Ala Asp Thr Gly Asn
        515                 520                 525
Ile Ile Ala Leu Glu Tyr Ser Glu Gln Tyr Pro Pro Ile Leu Ser Asn
    530                 535                 540
Phe Gly Met Gly Ser Lys Leu Ile Asn Tyr Tyr Arg Lys Glu Arg Pro
545                 550                 555                 560
Asn Asp Thr Ser Arg Pro Lys Ala Gln Ile Gly Glu Thr His Ile Leu
                565                 570                 575
Gly Val Glu Asp Arg Ser Pro Phe Trp Asn Phe Gly Glu Val Ala Pro
            580                 585                 590
Gly Asp Phe Val Pro Thr Leu Tyr Asn Asn Met Val Arg Ala Pro Ile
        595                 600                 605
Phe Lys His Asp Asn Lys Pro Thr Asp Phe Leu Leu Val Lys Ser Gln
    610                 615                 620
Gly Ala Gly Ser His Gln Lys Phe Tyr Leu Arg Gly Ile Asn Phe Asn
625                 630                 635                 640
Phe Ala Val Gly Asn Thr Phe Pro Val Glu Val Pro Ala Pro His Ser
                645                 650                 655
Arg Lys Val Thr Asn Ile Ser Lys Asn Arg Leu Lys Met Val Val Phe
            660                 665                 670
Arg Val Met Asn Ser Leu Gly Val Pro Arg Ile Ser Val Lys Asp Val
        675                 680                 685
Ser Lys His Phe Pro Glu His Ser Asp Met Gln Asn Arg Gln Arg Leu
    690                 695                 700
Lys Glu Phe Met Glu Tyr Gln Arg Gln Gly Glu Asp Gln Gly Tyr Trp
705                 710                 715                 720
Lys Val Arg Gly Leu Asn Asp Val Ile Pro Gly Glu Glu Glu Ile Arg
                725                 730                 735
Thr Met Ile Thr Pro Glu Asp Ser Ser Leu Met Asp Thr Met Gln Phe
            740                 745                 750
```

SUBSTITUTE SHEET (RULE 26)

```
Gly Gln Gln Val Leu Asp Asp Asn Met Val Leu Phe Gly Glu Gln Ser
        755                 760                 765
Arg Gln Glu Ser Ser Arg Ser Arg Lys Gly Asp Lys Arg Glu Asp Ser
    770                 775                 780
Ile Ala Asp Asp Ala Glu Asn Gly Asp Asp Ile Asn Lys Asp Lys Glu
785                 790                 795                 800
Lys Glu Val Glu Lys Glu Lys Glu Gln Glu Arg Glu Glu Glu Lys Gly
                805                 810                 815
Lys Asp Lys Glu Lys Asp Lys Asp Lys Glu Lys Asp Lys Thr Glu Lys
            820                 825                 830
Glu Lys Ser Lys Lys Ser Lys Glu Gln Asp Thr Glu Ile Asp Val Glu
        835                 840                 845
Glu Glu Leu Ala Pro Trp Asn Leu Ser Arg Asn Phe Val Ile Ala Asn
    850                 855                 860
Gln Thr Lys Thr Met Leu Gln Leu Asn Gly Glu Gly Asp Pro Thr Gly
865                 870                 875                 880
Ile Gly Leu Gly Phe Ser Met Leu Arg Ala Thr Gln Lys Asn Pro Phe
                885                 890                 895
Lys Pro Leu Phe Thr Pro Pro Pro Glu Asn Val Pro Lys Ser Asn Ala
            900                 905                 910
Ala Ala His Asn Gln Lys Leu Tyr Glu Gln Glu Ile Lys Arg Ile Trp
        915                 920                 925
Tyr Ser Gln Arg Ser Ser Leu Val Asp His Gly Glu Gly Thr Glu Ser
    930                 935                 940
Lys Leu Gln Gln Ile Tyr Asn Glu Tyr Pro Pro Ala Asp His Glu Leu
945                 950                 955                 960
Tyr Leu Lys Asn Lys Leu Glu Gln Asp Gln Gln Val Gln Gln Gln Gln
                965                 970                 975
Gln Asp Pro Ser Leu Gln Ala Asp Gln Gln Gln Gln Gln Gln Gln Gln
            980                 985                 990
Gln Asn Arg Val Leu Arg Ile Thr Arg Arg Val Arg Asp Glu Asn Gly
        995                 1000                1005
Ile Val His Arg Lys Val Glu Phe Ile His Asp Pro Arg Leu Ile Arg
    1010                1015                1020
Ala Tyr Val Lys Arg Lys Lys Gln Ile Glu Asp Glu Leu Leu Lys Asn
1025                1030                1035                104
Ala Asp Val Asp Glu Ile Leu Pro Thr Asn Asp Lys Glu Leu Asn Lys
                1045                1050                1055
Ile Arg Arg Lys Ala Leu Glu Glu Lys Leu Ala Asn Leu Glu Lys Arg
            1060                1065                1070
Ala Lys Gln Ser Arg Ala Lys Lys Pro Pro Lys Asp Leu Ile His Ala
        1075                1080                1085
Ala Ala Ala Ala Gly Ala Thr Ile Ile Asp Ala Asn Thr Val Met Leu
    1090                1095                1100
Pro Asp Gly Ser Tyr Val Ile Gly Gly Lys Gly Ile Gly Lys Gly Lys
1105                1110                1115                112
Ser Arg Thr Arg Arg Cys Lys Asn Cys Gly Ala Tyr Gly His Ile Arg
                1125                1130                1135
Thr Asn Ala Lys Cys Pro Leu Tyr Lys Lys Met Val Leu Gly Ile Asp
            1140                1145                1150
```

```
Asp Asp Ser Ala Ala Val Val Gly
        1155                1160

      <210> 3
      <211> 1066
      <212> PRT
      <213> S. cerevisiae

      <400> 3
Met Val Lys Gln Gln Gly Ser Gly Lys Thr Asn Leu Ala Asn Glu Asp
 1               5                  10                  15
Glu Ala Tyr Glu Ala Ile Phe Gly Gly Glu Phe Gly Ser Leu Glu Ile
            20                  25                  30
Gly Ser Tyr Ile Gly Gly Asp Glu Gly Ala Asn Ser Lys Asp Tyr Thr
        35                  40                  45
Glu His Leu Pro Asp Ala Val Asp Phe Glu Asp Glu Asp Glu Leu Ala
    50                  55                  60
Asp Asp Asp Asp Asp Leu Pro Glu Glu Ser Asp Ala Asn Leu His Pro
65                  70                  75                  80
Ala Met Met Thr Met Gly Ala Tyr Asp Asp Val Asn Glu Asn Gly Ala
                85                  90                  95
Val Leu Gly Ile Asp Ser Asn Ser Leu Asn Met Gln Leu Pro Glu Ile
            100                 105                 110
Asn Gly Asp Leu Ser Gln Gln Phe Ile Leu Glu Asp Asp Gly Gly Thr
        115                 120                 125
Pro Ala Thr Ser Asn Ala Leu Phe Met Gly Met Asp Ala Asn Glu Ile
    130                 135                 140
His Leu Ala Thr Glu Thr Gly Val Leu Asp Gly Ser Gly Ala Asn Glu
145                 150                 155                 160
Ile Gly His Ser Gln Leu Ser Ile Gly Gly Val Asn Gly Asn Asp Met
                165                 170                 175
Ser Ile Asn Gly Gly Phe Ile Met Glu Pro Asp Met Ser Asp Gly Lys
            180                 185                 190
His Lys Lys Ala Thr Lys Leu Asp Leu Ile Asn His Glu Lys Tyr Leu
        195                 200                 205
Leu Lys Lys Tyr Phe Pro Asp Phe Glu Lys Gly Lys Ile Leu Lys Trp
    210                 215                 220
Asn Lys Leu Ile Tyr Arg Arg Ser Val Pro Tyr His Trp His Ser Glu
225                 230                 235                 240
Ile Ser Arg Val Lys Lys Pro Phe Met Pro Leu Asn Leu Lys Phe Lys
                245                 250                 255
Val Gln Gln Asp Asp Lys Arg Leu Phe Asn Ser Arg Thr Ile Ser Tyr
            260                 265                 270
Val Ala Pro Ile Tyr Gln Gly Lys Asn Asn Leu Leu Gln Ser Asn Ser
        275                 280                 285
Ser Ala Ser Arg Arg Gly Leu Ile His Val Ser Ile Asp Glu Leu Phe
    290                 295                 300
Pro Ile Lys Glu Gln Gln Lys Lys Arg Lys Ile Ile His Asp Glu Lys
305                 310                 315                 320
Thr Ile Ser Glu Asp Leu Leu Ile Ala Thr Asp Asp Trp Asp Gln Glu
```

```
                  325                 330                 335
Lys Ile Ile Asn Gln Gly Thr Ser Ser Thr Ala Thr Leu Ala Asp Ser
              340                 345                 350
Ser Met Thr Pro Asn Leu Lys Phe Ser Gly Gly Tyr Lys Leu Lys Ser
          355                 360                 365
Leu Ile Glu Asp Val Ala Glu Asp Trp Gln Trp Asp Glu Asp Met Ile
      370                 375                 380
Ile Asp Ala Lys Leu Lys Glu Ser Lys His Ala Glu Leu Asn Met Asn
385                 390                 395                 400
Asp Glu Lys Leu Leu Leu Met Ile Glu Lys Thr Asn Asn Leu Ala Gln
                  405                 410                 415
Gln Lys Gln Gln Leu Asp Ser Ser Asn Leu Ile Leu Pro Leu Asn Glu
              420                 425                 430
Thr Ile Leu Gln Gln Lys Phe Asn Leu Ser Asn Asp Asp Lys Tyr Gln
          435                 440                 445
Ile Leu Lys Lys Thr His Gln Thr Lys Val Arg Ser Thr Ile Ser Asn
      450                 455                 460
Leu Asn Ile Gln His Ser Gln Pro Ala Ile Asn Leu Gln Ser Pro Phe
465                 470                 475                 480
Tyr Lys Val Ala Val Pro Arg Tyr Gln Leu Arg His Phe His Arg Glu
                  485                 490                 495
Asn Phe Gly Ser His Ile Arg Pro Gly Thr Lys Ile Val Phe Ser Lys
              500                 505                 510
Leu Lys Ala Arg Lys Arg Lys Arg Asp Lys Gly Lys Asp Val Lys Glu
          515                 520                 525
Ser Phe Ser Thr Ser Gln Asp Leu Thr Ile Gly Asp Thr Ala Pro Val
      530                 535                 540
Tyr Leu Met Glu Tyr Ser Glu Gln Thr Pro Val Ala Leu Ser Lys Phe
545                 550                 555                 560
Gly Met Ala Asn Lys Leu Ile Asn Tyr Tyr Arg Lys Ala Asn Glu Gln
                  565                 570                 575
Asp Thr Leu Arg Pro Lys Leu Pro Val Gly Glu Thr His Val Leu Gly
              580                 585                 590
Val Gln Asp Lys Ser Pro Phe Trp Asn Phe Gly Phe Val Glu Pro Gly
          595                 600                 605
His Ile Val Pro Thr Leu Tyr Asn Asn Met Ile Arg Ala Pro Val Phe
      610                 615                 620
Lys His Asp Ile Ser Gly Thr Asp Phe Leu Leu Thr Lys Ser Ser Gly
625                 630                 635                 640
Phe Gly Ile Ser Asn Arg Phe Tyr Leu Arg Asn Ile Asn His Leu Phe
                  645                 650                 655
Thr Val Gly Gln Thr Phe Pro Val Glu Glu Ile Pro Gly Pro Asn Ser
              660                 665                 670
Arg Lys Val Thr Ser Met Lys Ala Thr Arg Leu Lys Met Ile Ile Tyr
          675                 680                 685
Arg Ile Leu Asn His Asn His Ser Lys Ala Ile Ser Ile Asp Pro Ile
      690                 695                 700
Ala Lys His Phe Pro Asp Gln Asp Tyr Gly Gln Asn Arg Gln Lys Val
705                 710                 715                 720
Lys Glu Phe Met Lys Tyr Gln Arg Asp Gly Pro Glu Lys Gly Leu Trp
```

09940199 1.1990

```
                725                 730                 735
Arg Leu Lys Asp Asp Glu Lys Leu Leu Asp Asn Glu Ala Val Lys Ser
            740                 745                 750
Leu Ile Thr Pro Glu Gln Ile Ser Gln Val Glu Ser Met Ser Gln Gly
        755                 760                 765
Leu Gln Phe Gln Glu Asp Asn Glu Ala Tyr Asn Phe Asp Ser Lys Leu
    770                 775                 780
Lys Ser Leu Glu Glu Asn Leu Leu Pro Trp Asn Ile Thr Lys Asn Phe
785                 790                 795                 800
Ile Asn Ser Thr Gln Met Arg Ala Met Ile Gln Ile His Gly Val Gly
                805                 810                 815
Asp Pro Thr Gly Cys Gly Glu Gly Phe Ser Phe Leu Lys Thr Ser Met
            820                 825                 830
Lys Gly Gly Phe Val Lys Ser Gly Ser Pro Ser Ser Asn Asn Asn Ser
        835                 840                 845
Ser Asn Lys Lys Gly Thr Asn Thr His Ser Tyr Asn Val Ala Gln Gln
    850                 855                 860
Gln Lys Ala Tyr Asp Glu Glu Ile Ala Lys Thr Trp Tyr Thr His Thr
865                 870                 875                 880
Lys Ser Leu Ser Ile Ser Asn Pro Phe Glu Glu Met Thr Asn Pro Asp
                885                 890                 895
Glu Ile Asn Gln Thr Asn Lys His Val Lys Thr Asp Arg Asp Asp Lys
            900                 905                 910
Lys Ile Leu Lys Ile Val Arg Lys Lys Arg Asp Glu Asn Gly Ile Ile
        915                 920                 925
Gln Arg Gln Thr Ile Phe Ile Arg Asp Pro Arg Val Ile Gln Gly Tyr
    930                 935                 940
Ile Lys Ile Lys Glu Gln Asp Lys Glu Asp Val Asn Lys Leu Leu Glu
945                 950                 955                 960
Glu Asp Thr Ser Lys Ile Asn Asn Leu Glu Glu Leu Glu Lys Gln Lys
                965                 970                 975
Lys Leu Leu Gln Leu Glu Leu Ala Asn Leu Glu Lys Ser Gln Gln Arg
            980                 985                 990
Arg Ala Ala Arg Gln Asn Ser Lys Arg Asn Gly Gly Ala Thr Arg Thr
        995                1000                1005
Glu Asn Ser Val Asp Asn Gly Ser Asp Leu Ala Gly Val Thr Asp Gly
    1010                1015                1020
Lys Ala Ala Arg Asn Lys Gly Lys Asn Thr Thr Arg Arg Cys Ala Thr
1025                1030                1035                104
Cys Gly Gln Ile Gly His Ile Arg Thr Asn Lys Ser Cys Pro Met Tyr
                1045                1050                1055
Ser Ser Lys Asp Asn Pro Ala Ser Pro Lys
            1060                1065
```

<210> 4
<211> 979
<212> PRT
<213> S. pombe

<400> 4

```
Met Ser Phe Asp Gly Leu Ile Val Glu Asn Glu Asn Thr Lys Ser Gly
 1               5                  10                  15
Tyr Asn Asp Gly Asn Asp Leu Thr Asp Leu Phe Lys Gln Asn Gly Thr
            20                  25                  30
Asp Met Ser Val Ile Asn Ser Leu Leu Gly Asp Thr Asn Asn Pro Gly
        35                  40                  45
Met Asn Glu Ser Pro Lys Ile Leu Asp Ser Ser Phe Glu Asn Ser Asn
    50                  55                  60
Pro Gln Asp Gly Pro Asn Tyr Glu Asp Phe Asp Phe Met Gly Ser Ile
65                  70                  75                  80
His Lys Glu Phe Gly Asn Asn Ile Asn Glu Met Asp Asp Met Glu Asp
                85                  90                  95
Val Ser Asp Asp Asn Leu Pro Glu Glu Glu Gln Ala Val Asn Tyr Thr
            100                 105                 110
Gly Asp Lys Asp Asp Glu Asp Phe Gly Lys Leu Leu Ala Lys Glu Met
        115                 120                 125
Gly Glu Glu Ala Ala Gly Gln Val Leu Ser Gly Val Gly Phe Ser Ile
    130                 135                 140
Pro Ser Gly Leu Val Pro Pro Ser Glu Pro Ser Lys Thr Val Ser Ser
145                 150                 155                 160
Thr Thr Glu Glu Leu Gln Asn Glu Ala Gln Ile Arg Glu Ser Ile Val
                165                 170                 175
Lys Thr Phe Phe Pro Thr Phe Glu Arg Gly Val Leu Leu Asn Phe Ser
            180                 185                 190
Glu Leu Phe Lys Pro Lys Pro Val Lys Leu Ala Pro Pro Lys Lys Lys
        195                 200                 205
Thr Pro Lys Val Cys Val Pro Gly Arg Leu Thr Leu Glu Val Asp Thr
    210                 215                 220
Asp Tyr Ala Ile Ile Phe Asn Ser Lys Lys Ser Leu Pro Leu Lys Arg
225                 230                 235                 240
Asn Val Val Ser Pro Ile Ser Thr His Thr Lys Lys Arg Arg Arg Thr
                245                 250                 255
Ala Asn Thr Ser Gln Arg Asn Asp Gly Leu Asp Leu Asn Thr Val Phe
            260                 265                 270
Thr Thr Asn Asp Trp Glu Lys Asn Ile Ile Tyr Asp Glu Ser Asp Val
        275                 280                 285
Asn Lys Thr Asn Gln Ser Ser Phe Phe Ile Asp Lys Ser Leu Val Asp
    290                 295                 300
Ile Asp Phe Ala Phe Asp Glu Asn Ile Phe Asp Gly Asp Thr Gly Thr
305                 310                 315                 320
Ser Lys Val Val Leu Asn Leu Asn Asp Pro Lys Leu Leu Leu Gln Pro
                325                 330                 335
Gln Leu Pro Lys Lys Glu Asp Ser Gln Arg Ser Phe Ala Asp Thr His
            340                 345                 350
Gln Arg Asn Ser Leu Ala Trp Lys Phe Asn Ile Ser Asn Asp Pro Ala
        355                 360                 365
Tyr Glu Met Leu Lys Gln Asn His Gln Ser Lys Val Arg Asn Thr Leu
    370                 375                 380
Ser Gln Leu Ala Ile Glu His Ala Ala Phe Ala Glu Lys Leu Thr Phe
385                 390                 395                 400
```

```
Pro Tyr Tyr Lys Thr Arg Leu Ser Lys Arg Ala Val Arg Ser Tyr His
                405                 410                 415
Arg Pro Thr Met Ser Phe Lys Pro Asn Ala Ala Ile Val Phe Ser Pro
            420                 425                 430
Leu Ile Val Arg Lys Arg Ser Lys Asp Lys His Lys Ser Glu Arg Glu
        435                 440                 445
Leu Ile Pro Thr Thr Lys Glu Ile Thr Met Gly Asp Thr Thr His Ala
    450                 455                 460
Ile Leu Val Glu Phe Ser Glu Glu His Pro Ala Val Leu Ser Asn Ala
465                 470                 475                 480
Gly Met Ala Ser Arg Ile Val Asn Tyr Tyr Arg Lys Lys Asn Glu Gln
                485                 490                 495
Asp Glu Ser Arg Pro Lys Leu Glu Val Gly Glu Ser His Val Leu Asp
            500                 505                 510
Val Gln Asp Arg Ser Pro Phe Trp Asn Phe Gly Ser Val Glu Pro Gly
        515                 520                 525
Glu Ile Thr Pro Thr Leu Tyr Asn Lys Met Ile Arg Ala Pro Leu Phe
    530                 535                 540
Lys His Glu Val Pro Pro Thr Asp Phe Ile Leu Ile Arg Asn Ser Ser
545                 550                 555                 560
Ser Tyr Gly Ser Lys Tyr Tyr Leu Lys Asn Ile Asn His Met Phe Val
                565                 570                 575
Ser Gly Gln Thr Phe Pro Val Thr Asp Val Pro Gly Pro His Ser Arg
            580                 585                 590
Lys Val Thr Thr Ala Ser Lys Asn Arg Leu Lys Met Leu Val Phe Arg
        595                 600                 605
Leu Ile Arg Arg Ser Pro Asn Gly Gly Leu Phe Ile Arg Gln Leu Ser
    610                 615                 620
Lys His Phe Ser Asp Gln Asn Glu Met Gln Ile Arg Gln Arg Leu Lys
625                 630                 635                 640
Glu Phe Met Glu Tyr Lys Lys Lys Gly Asp Gly Pro Gly Tyr Trp Lys
                645                 650                 655
Leu Lys Ser Asn Glu Val Val Pro Asp Glu Ala Gly Thr Arg Ser Met
            660                 665                 670
Val Ser Pro Glu Thr Val Cys Leu Leu Glu Ser Met Gln Val Gly Val
        675                 680                 685
Arg Gln Leu Glu Asp Ala Gly Tyr Gly Lys Thr Met Asp Glu Ile Asn
    690                 695                 700
Asp Asp Glu Asp Glu Glu Gln Pro Ala Glu Gln Leu Leu Ala Pro Trp
705                 710                 715                 720
Ile Thr Thr Arg Asn Phe Ile Asn Ala Thr Gln Gly Lys Ala Met Leu
                725                 730                 735
Thr Leu Phe Gly Glu Gly Asp Pro Thr Gly Ile Gly Glu Gly Tyr Ser
            740                 745                 750
Phe Ile Arg Thr Ser Met Lys Gly Gly Phe Lys Pro Ala Gly Glu Thr
        755                 760                 765
Ala Asp Asp Lys Pro Glu Pro Gln Thr Lys Asn Ala His Ala Tyr Asn
    770                 775                 780
Val Ala Lys Gln Gln Arg Ala Tyr Glu Glu Glu Ile Asn Arg Ile Trp
785                 790                 795                 800
```

```
Asn Ala Gln Lys Arg Gly Leu Ser Ile Asn Asn Leu Glu Glu Leu Ala
                805                 810                 815
Lys Lys Tyr Gly Ile Asn Ser Ile His Asp Asp Tyr Val Glu Ser Asn
            820                 825                 830
Glu Glu Thr Thr Arg Glu Glu Thr Pro Ser Ser Asp Lys Val Leu Arg
        835                 840                 845
Ile Val Arg Leu Tyr Arg Asp Lys Asn Gly Asn Leu Glu Arg Lys Gln
    850                 855                 860
Glu Thr Ile His Asp Pro Ile Val Ile His Ala Tyr Leu Lys Lys Arg
865                 870                 875                 880
Arg Glu Ile Asp Glu Gln Ser Thr Ala Leu Asp Ala Val Val Pro Thr
                885                 890                 895
Gly Asp Glu Ala Ile Asp Arg Arg Asn Arg Arg Arg Leu Glu Gln Glu
            900                 905                 910
Leu Ala Lys Ser Gln Lys Asn Trp Glu Arg Arg Arg Ala Arg His Ala
        915                 920                 925
Ala Lys Glu Gly Ile Asn Leu Asn Gly Glu Gly Arg Lys Pro Thr Thr
    930                 935                 940
Arg Lys Cys Ser Asn Cys Gly Gln Val Gly His Met Lys Thr Asn Lys
945                 950                 955                 960
Ile Cys Pro Leu Phe Gly Arg Pro Glu Gly Gly Leu Ala Thr Met Leu
                965                 970                 975
Asp Lys Asn
```

<210> 5
<211> 2066
<212> PRT
<213> Drosophila

<400> 5

```
Met Glu Ser Asp Asn Ser Asp Asp Glu Gly Ser Ile Gly Asn Gly Leu
 1               5                  10                  15
Asp Leu Thr Gly Ile Leu Phe Gly Asn Ile Asp Ser Glu Gly Arg Leu
            20                  25                  30
Leu Gln Asp Asp Asp Gly Glu Gly Arg Gly Gly Thr Gly Phe Asp Ala
        35                  40                  45
Glu Leu Arg Glu Asn Ile Gly Ser Leu Ser Lys Leu Gly Leu Asp Ser
    50                  55                  60
Met Leu Leu Glu Val Ile Asp Leu Lys Glu Ala Glu Pro Pro Ser Asp
65                  70                  75                  80
Asp Glu Glu Glu Glu Asp Ala Arg Pro Ser Ala Val Ser Ala Ser Gly
                85                  90                  95
Gly Met Ser Ala Phe Asp Ala Leu Lys Ala Gly Val Lys Arg Glu Glu
            100                 105                 110
Arg Glu Asp Gly Ala Val Lys Ala Gln Asp Asp Ala Ile Asp Tyr Ser
        115                 120                 125
Asp Ile Thr Glu Leu Ser Glu Asp Cys Pro Arg Thr Pro Pro Glu Glu
    130                 135                 140
Thr Ser Thr Tyr Asp Asp Leu Glu Asp Ala Ile Pro Ala Ser Lys Val
```

```
145                 150                 155                 160
Glu Ala Lys Leu Thr Lys Asp Asp Lys Glu Leu Met Pro Pro Pro Ser
                165                 170                 175
Ala Pro Met Arg Ser Gly Ser Gly Gly Gly Ile Glu Glu Pro Ala Lys
            180                 185                 190
Ser Asn Asp Ala Ser Ser Pro Ser Asp Asp Ser Lys Ser Thr Asp Ser
        195                 200                 205
Lys Asp Ala Asp Arg Lys Leu Asp Thr Pro Leu Ala Asp Ile Leu Pro
    210                 215                 220
Ser Lys Tyr Gln Asn Val Asp Val Arg Glu Leu Phe Pro Asp Phe Arg
225                 230                 235                 240
Pro Gln Lys Val Leu Arg Phe Ser Arg Leu Phe Gly Pro Gly Lys Pro
                245                 250                 255
Thr Ser Leu Pro Gln Ile Trp Arg His Val Arg Lys Arg Arg Arg Lys
            260                 265                 270
Arg Asn Gln Ser Arg Asp Gln Lys Thr Thr Asn Thr Gly Gly Ser Asp
        275                 280                 285
Ser Pro Ser Asp Thr Glu Glu Pro Arg Lys Arg Gly Phe Ser Leu His
    290                 295                 300
Tyr Ala Ala Glu Pro Thr Pro Ala Glu Cys Met Ser Asp Asp Glu Asp
305                 310                 315                 320
Lys Leu Leu Gly Asp Phe Asn Ser Glu Asp Val Arg Pro Glu Gly Pro
                325                 330                 335
Asp Asn Gly Glu Asn Ser Asp Gln Lys Pro Lys Val Ala Asp Trp Arg
            340                 345                 350
Phe Gly Pro Ala Gln Ile Trp Tyr Asp Ile Leu Glu Val Pro Asp Ser
        355                 360                 365
Gly Glu Gly Phe Asn Tyr Gly Phe Lys Thr Lys Ala Ala Ser Thr Ser
    370                 375                 380
Ser Gln Gln Gln Leu Lys Asp Glu Arg Arg Val Lys Ser Pro Glu Asp
385                 390                 395                 400
Asp Val Glu Asp Pro Ser Ile Ala Asp Asp Ala Phe Leu Met Val Ser
                405                 410                 415
Gln Leu His Trp Glu Asp Asp Val Val Trp Asp Gly Asn Asp Ile Lys
            420                 425                 430
Ala Lys Val Leu Gln Lys Leu Asn Ser Lys Thr Asn Ala Ala Gly Trp
        435                 440                 445
Leu Pro Ser Ser Gly Ser Arg Thr Ala Gly Ala Phe Ser Gln Pro Gly
    450                 455                 460
Lys Pro Ser Met Pro Val Gly Ser Gly Ser Ser Lys Gln Gly Ser Gly
465                 470                 475                 480
Ala Ser Ser Lys Lys Ala Gln Gln Asn Ala Gln Ala Lys Pro Ala Glu
                485                 490                 495
Ala Pro Asp Asp Thr Trp Tyr Ser Leu Phe Pro Val Glu Asn Glu Glu
            500                 505                 510
Leu Ile Tyr Tyr Lys Trp Glu Asp Glu Val Ile Trp Asp Ala Gln Gln
        515                 520                 525
Val Ser Lys Val Pro Lys Pro Lys Val Leu Thr Leu Asp Pro Asn Asp
    530                 535                 540
Glu Asn Ile Ile Leu Gly Ile Pro Asp Asp Ile Asp Pro Ser Lys Ile
```

09940199 A1 991028

```
545                 550                 555                 560
Asn Lys Ser Thr Gly Pro Pro Pro Lys Ile Lys Ile Pro His Pro His
                565                 570                 575
Val Lys Lys Ser Lys Ile Leu Leu Gly Lys Ala Gly Val Ile Asn Val
            580                 585                 590
Leu Ala Glu Asp Thr Pro Pro Pro Pro Pro Lys Ser Pro Asp Arg Asp
        595                 600                 605
Pro Phe Asn Ile Ser Asn Asp Thr Tyr Tyr Thr Pro Lys Thr Glu Pro
    610                 615                 620
Thr Leu Arg Leu Lys Val Gly Gly Asn Leu Ile Gln His Ser Thr Pro
625                 630                 635                 640
Val Val Glu Leu Arg Ala Pro Phe Val Pro Thr His Met Gly Pro Met
                645                 650                 655
Asn Val Arg Ala Phe His Arg Pro Pro Leu Lys Lys Tyr Ser His Gly
            660                 665                 670
Pro Met Ala Gln Ser Ile Pro His Pro Val Phe Pro Leu Leu Lys Thr
        675                 680                 685
Ile Ala Lys Lys Ala Lys Gln Arg Glu Val Glu Arg Ile Ala Ser Gly
    690                 695                 700
Gly Gly Asp Val Phe Phe Met Arg Asn Pro Glu Asp Leu Ser Gly Arg
705                 710                 715                 720
Asp Gly Asp Ile Val Leu Ala Glu Phe Cys Glu Glu His Pro Pro Leu
                725                 730                 735
Ile Asn Gln Val Gly Met Cys Ser Lys Ile Lys Asn Tyr Tyr Lys Arg
            740                 745                 750
Lys Ala Glu Lys Asp Ser Gly Pro Gln Asp Tyr Val Tyr Gly Glu Val
        755                 760                 765
Ala Phe Ala His Thr Ser Pro Phe Leu Gly Ile Leu His Pro Gly Gln
    770                 775                 780
Cys Ile Gln Ala Ile Glu Asn Asn Met Tyr Arg Ala Pro Ile Tyr Pro
785                 790                 795                 800
His Lys Met Ala His Asn Asp Phe Leu Val Ile Arg Thr Arg Asn Asn
                805                 810                 815
Tyr Trp Ile Arg Ser Val Asn Ser Ile Tyr Thr Val Gly Gln Glu Cys
            820                 825                 830
Pro Leu Tyr Glu Val Pro Gly Pro Asn Ser Lys Arg Ala Asn Asn Phe
        835                 840                 845
Thr Arg Asp Phe Leu Gln Val Phe Ile Tyr Arg Leu Phe Trp Lys Ser
    850                 855                 860
Arg Asp Asn Pro Arg Arg Ile Arg Met Asp Asp Ile Lys Gln Ala Phe
865                 870                 875                 880
Pro Ala His Ser Glu Ser Ser Ile Arg Lys Arg Leu Lys Gln Cys Ala
                885                 890                 895
Asp Phe Lys Arg Thr Gly Met Asp Ser Asn Trp Trp Val Ile Lys Pro
            900                 905                 910
Glu Phe Arg Leu Pro Ser Glu Glu Glu Ile Arg Ala Met Val Ser Pro
        915                 920                 925
Glu Gln Cys Cys Ala Tyr Phe Ser Met Ile Ala Ala Glu Gln Arg Leu
    930                 935                 940
Lys Asp Ala Gly Tyr Gly Glu Lys Phe Leu Phe Ala Pro Gln Glu Asp
```

```
945                 950                 955                 960
Asp Asp Glu Glu Ala Gln Leu Lys Leu Asp Asp Glu Val Lys Val Ala
                965                 970                 975
Pro Trp Asn Thr Thr Arg Ala Tyr Ile Gln Ala Met Arg Gly Lys Cys
            980                 985                 990
Leu Leu Gln Leu Ser Gly Pro Ala Asp Pro Thr Gly Cys Gly Glu Gly
        995                 1000                1005
Phe Ser Tyr Val Arg Val Pro Asn Lys Pro Thr Gln Thr Lys Glu Glu
    1010                1015                1020
Gln Glu Ser Gln Pro Lys Arg Ser Val Thr Gly Thr Asp Ala Asp Leu
1025                1030                1035                104
Arg Arg Leu Pro Leu Gln Arg Ala Lys Glu Leu Leu Arg Gln Phe Lys
                1045                1050                1055
Val Pro Glu Glu Glu Ile Lys Lys Leu Ser Arg Trp Glu Val Ile Asp
            1060                1065                1070
Val Val Arg Thr Leu Ser Thr Glu Lys Ala Lys Ala Gly Glu Glu Gly
        1075                1080                1085
Met Asp Lys Phe Ser Arg Gly Asn Arg Phe Ser Ile Ala Glu His Gln
    1090                1095                1100
Glu Arg Tyr Lys Glu Glu Cys Gln Arg Ile Phe Asp Leu Gln Asn Arg
1105                1110                1115                112
Val Leu Ala Ser Ser Glu Val Leu Ser Thr Asp Glu Ala Glu Ser Ser
                1125                1130                1135
Ala Ser Glu Glu Ser Asp Leu Glu Glu Leu Gly Lys Asn Leu Glu Asn
            1140                1145                1150
Met Leu Ser Asn Lys Lys Thr Ser Thr Gln Leu Ser Arg Glu Arg Glu
        1155                1160                1165
Glu Leu Glu Arg Gln Glu Leu Leu Arg Gln Leu Asp Glu Glu His Gly
    1170                1175                1180
Gly Pro Ser Gly Ser Gly Gly Ala Lys Gly Ala Lys Gly Lys Asp Asp
1185                1190                1195                120
Pro Gly Gln Gln Met Leu Ala Thr Asn Asn Gln Gly Arg Ile Leu Arg
                1205                1210                1215
Ile Thr Arg Thr Phe Arg Gly Asn Asp Gly Lys Glu Tyr Thr Arg Val
            1220                1225                1230
Glu Thr Val Arg Arg Gln Pro Val Ile Asp Ala Tyr Ile Lys Ile Arg
        1235                1240                1245
Thr Thr Lys Asp Glu Gln Phe Ile Lys Gln Phe Ala Thr Leu Asp Glu
    1250                1255                1260
Gln Gln Lys Glu Glu Met Lys Arg Glu Lys Arg Arg Ile Gln Glu Gln
1265                1270                1275                128
Leu Arg Arg Ile Lys Arg Asn Gln Glu Arg Glu Arg Leu Ala Gln Leu
                1285                1290                1295
Ala Gln Asn Gln Lys Leu Gln Pro Gly Gly Met Pro Thr Ser Leu Gly
            1300                1305                1310
Asp Pro Lys Ser Ser Gly Gly His Ser His Lys Glu Arg Asp Ser Gly
        1315                1320                1325
Tyr Lys Glu Val Ser Pro Ser Arg Lys Lys Phe Lys Leu Lys Pro Asp
    1330                1335                1340
Leu Lys Leu Lys Cys Gly Ala Cys Gly Gln Val Gly His Met Arg Thr
```

1345 1350 1355 136
Asn Lys Ala Cys Pro Leu Tyr Ser Gly Met Gln Ser Ser Leu Ser Gln
1365 1370 1375
Ser Asn Pro Ser Leu Ala Asp Asp Phe Asp Glu Gln Ser Glu Lys Glu
1380 1385 1390
Met Thr Met Asp Asp Asp Asp Leu Val Asn Val Asp Gly Thr Lys Val
1395 1400 1405
Thr Leu Ser Ser Lys Ile Leu Lys Arg His Gly Gly Asp Asp Gly Lys
1410 1415 1420
Arg Arg Ser Gly Ser Ser Ser Gly Phe Thr Leu Lys Val Pro Arg Asp
1425 1430 1435 144
Ala Met Gly Lys Lys Lys Arg Arg Val Gly Gly Asp Leu His Cys Asp
1445 1450 1455
Tyr Leu Gln Arg His Asn Lys Thr Ala Asn Arg Arg Arg Thr Asp Pro
1460 1465 1470
Val Val Val Leu Ser Ser Ile Leu Glu Ile Ile His Asn Glu Leu Arg
1475 1480 1485
Ser Met Pro Asp Val Ser Pro Phe Leu Phe Pro Val Ser Ala Lys Lys
1490 1495 1500
Val Pro Asp Tyr Tyr Arg Val Val Thr Lys Pro Met Asp Leu Gln Thr
1505 1510 1515 152
Met Arg Glu Tyr Ile Arg Gln Arg Arg Tyr Thr Ser Arg Glu Met Phe
1525 1530 1535
Leu Glu Asp Leu Lys Gln Ile Val Asp Asn Ser Leu Ile Tyr Asn Gly
1540 1545 1550
Pro Gln Ser Ala Tyr Thr Leu Ala Ala Gln Arg Met Phe Ser Ser Cys
1555 1560 1565
Phe Glu Leu Leu Ala Glu Arg Glu Asp Lys Leu Met Arg Leu Glu Lys
1570 1575 1580
Ala Ile Asn Pro Leu Leu Asp Asp Asp Asp Gln Val Ala Leu Ser Phe
1585 1590 1595 160
Ile Phe Asp Lys Leu His Ser Gln Ile Lys Gln Leu Pro Glu Ser Trp
1605 1610 1615
Pro Phe Leu Lys Pro Val Asn Lys Lys Gln Val Lys Asp Tyr Tyr Thr
1620 1625 1630
Val Ile Lys Arg Pro Met Asp Leu Glu Thr Ile Gly Lys Asn Ile Glu
1635 1640 1645
Ala His Arg Tyr His Ser Arg Ala Glu Tyr Leu Ala Asp Ile Glu Leu
1650 1655 1660
Ile Ala Thr Asn Cys Glu Gln Tyr Asn Gly Ser Asp Thr Arg Tyr Thr
1665 1670 1675 168
Lys Phe Ser Lys Lys Ile Leu Glu Tyr Ala Gln Thr Gln Leu Ile Glu
1685 1690 1695
Phe Ser Glu His Cys Gly Gln Leu Glu Asn Asn Ile Ala Lys Thr Gln
1700 1705 1710
Glu Arg Ala Arg Glu Asn Ala Pro Glu Phe Asp Glu Ala Trp Gly Asn
1715 1720 1725
Asp Asp Tyr Asn Phe Asp Arg Gly Ser Arg Ala Ser Ser Pro Gly Asp
1730 1735 1740
Asp Tyr Ile Asp Val Glu Gly His Gly Gly His Ala Ser Ser Ser Asn

1745 1750 1755 176
Ser Ile His Arg Ser Met Gly Ala Glu Ala Gly Ser Ser His Thr Ala
1765 1770 1775
Pro Ala Val Arg Lys Pro Ala Pro Pro Gly Pro Gly Glu Val Lys Arg
1780 1785 1790
Gly Arg Gly Arg Pro Arg Lys Gln Arg Asp Pro Val Glu Glu Asp Leu
1795 1800 1805
Gln Cys Ser Thr Asp Asp Glu Asp Asp Asp Glu Glu Glu Asp Phe Gln
1810 1815 1820
Glu Val Ser Glu Asp Glu Asn Asn Ala Ala Ser Ile Leu Asp Gln Gly
1825 1830 1835 184
Glu Arg Ile Asn Ala Pro Ala Asp Ala Met Asp Gly Met Phe Asp Pro
1845 1850 1855
Lys Asn Ile Lys Thr Glu Ile Asp Leu Glu Ala His Gln Met Ala Asp
1860 1865 1870
Glu Ser Met Asp Val Asp Pro Asn Tyr Asp Pro Ser Asp Phe Leu Ala
1875 1880 1885
Met His Lys Gln Arg Gln Ser Leu Gly Glu Pro Ser Ser Leu Gln Gly
1890 1895 1900
Ala Phe Thr Asn Phe Leu Ser His Glu Gln Asp Asp Asn Gly Pro Tyr
1905 1910 1915 192
Asn Pro Ala Glu Ala Ser Thr Ser Ala Ala Ser Gly Ala Asp Leu Gly
1925 1930 1935
Met Asp Ala Ser Met Ala Met Gln Met Ala Pro Glu Met Pro Val Asn
1940 1945 1950
Thr Met Asn Asn Gly Met Gly Ile Asp Asp Asp Leu Asp Ile Ser Glu
1955 1960 1965
Ser Asp Glu Glu Asp Asp Gly Ser Arg Val Arg Ile Lys Lys Glu Val
1970 1975 1980
Phe Asp Asp Gly Asp Tyr Ala Leu Gln His Gln Gln Met Gly Gln Ala
1985 1990 1995 200
Ala Ser Gln Ser Gln Ile Tyr Met Val Asp Ser Ser Asn Glu Pro Thr
2005 2010 2015
Thr Leu Asp Tyr Gln Gln Pro Pro Gln Leu Asp Phe Gln Gln Val Gln
2020 2025 2030
Glu Met Glu Gln Leu Gln His Gln Val Met Pro Pro Met Gln Ser Glu
2035 2040 2045
Gln Leu Gln Gln Gln Gln Thr Pro Gln Gly Asp Asn Asp Tyr Ala Trp
2050 2055 2060
Thr Phe
2065

<210> 6
<211> 1872
<212> PRT
<213> Human

<400> 6
Met Gly Pro Gly Cys Asp Leu Leu Leu Arg Thr Ala Ala Thr Ile Thr
1 5 10 15

```
Ala Ala Ala Ile Met Ser Asp Thr Asp Ser Asp Glu Asp Ser Ala Gly
                20                  25                  30
Gly Gly Pro Phe Ser Leu Ala Gly Phe Leu Phe Gly Asn Ile Asn Gly
            35                  40                  45
Ala Gly Gln Leu Glu Gly Glu Ser Val Leu Asp Asp Glu Cys Lys Lys
        50                  55                  60
His Leu Ala Gly Leu Gly Ala Leu Gly Leu Gly Ser Leu Ile Thr Glu
65                  70                  75                  80
Leu Thr Ala Asn Glu Glu Leu Thr Gly Thr Asp Gly Ala Leu Val Asn
                85                  90                  95
Asp Glu Gly Trp Val Arg Ser Thr Glu Asp Ala Val Asp Tyr Ser Asp
            100                 105                 110
Ile Asn Glu Val Ala Glu Asp Glu Ser Arg Arg Tyr Gln Gln Thr Met
        115                 120                 125
Gly Ser Leu Gln Pro Leu Cys His Ser Asp Tyr Asp Glu Asp Asp Tyr
    130                 135                 140
Asp Ala Asp Cys Glu Asp Ile Asp Cys Lys Leu Met Pro Pro Pro Pro
145                 150                 155                 160
Pro Pro Pro Gly Pro Met Lys Lys Asp Lys Asp Gln Asp Ser Ile Thr
                165                 170                 175
Gly Glu Lys Val Asp Phe Ser Ser Ser Ser Asp Ser Glu Ser Glu Met
            180                 185                 190
Gly Pro Gln Glu Ala Thr Gln Ala Glu Ser Glu Asp Gly Lys Leu Thr
        195                 200                 205
Leu Pro Leu Ala Gly Ile Met Gln His Asp Ala Thr Lys Leu Leu Pro
    210                 215                 220
Ser Val Thr Glu Leu Phe Pro Glu Phe Arg Pro Gly Lys Val Leu Arg
225                 230                 235                 240
Phe Leu Arg Leu Phe Gly Pro Gly Lys Asn Val Pro Ser Val Trp Arg
                245                 250                 255
Ser Ala Arg Arg Lys Arg Lys Lys Lys His Arg Glu Leu Ile Gln Glu
            260                 265                 270
Glu Gln Ile Gln Glu Val Glu Cys Ser Val Glu Ser Glu Val Ser Gln
        275                 280                 285
Lys Ser Leu Trp Asn Tyr Asp Tyr Ala Pro Pro Pro Pro Pro Glu Gln
    290                 295                 300
Cys Leu Ser Asp Asp Glu Ile Thr Met Met Ala Pro Val Glu Ser Lys
305                 310                 315                 320
Phe Ser Gln Ser Thr Gly Asp Ile Asp Lys Val Thr Asp Thr Lys Pro
                325                 330                 335
Arg Val Ala Glu Trp Arg Tyr Gly Pro Ala Arg Leu Trp Tyr Asp Met
            340                 345                 350
Leu Gly Val Pro Glu Asp Gly Ser Gly Phe Asp Tyr Gly Phe Lys Leu
        355                 360                 365
Arg Lys Thr Glu His Glu Pro Val Ile Lys Ser Arg Met Ile Glu Glu
    370                 375                 380
Phe Arg Lys Leu Glu Glu Asn Asn Gly Thr Asp Leu Leu Ala Asp Glu
385                 390                 395                 400
Asn Phe Leu Met Val Thr Gln Leu His Trp Glu Asp Asp Ile Ile Trp
                405                 410                 415
```

```
Asp Gly Glu Asp Val Lys His Lys Gly Thr Lys Pro Gln Arg Ala Ser
            420                 425                 430
Leu Ala Gly Trp Leu Pro Ser Ser Met Thr Arg Asn Ala Met Ala Tyr
        435                 440                 445
Asn Val Gln Gln Gly Phe Ala Ala Thr Leu Asp Asp Asp Lys Pro Trp
    450                 455                 460
Tyr Ser Ile Phe Pro Ile Asp Asn Glu Asp Leu Val Tyr Gly Arg Trp
465                 470                 475                 480
Glu Asp Asn Ile Ile Trp Asp Ala Gln Ala Met Pro Arg Leu Leu Glu
                485                 490                 495
Pro Pro Val Leu Thr Leu Asp Pro Asn Asp Glu Asn Leu Ile Leu Glu
            500                 505                 510
Ile Pro Asp Glu Lys Glu Glu Ala Thr Ser Asn Ser Pro Ser Lys Glu
        515                 520                 525
Ser Lys Lys Glu Ser Ser Leu Lys Lys Ser Arg Ile Leu Leu Gly Lys
    530                 535                 540
Thr Gly Val Ile Lys Glu Glu Pro Gln Gln Asn Met Ser Gln Pro Glu
545                 550                 555                 560
Val Lys Asp Pro Trp Asn Leu Ser Asn Asp Glu Tyr Tyr Tyr Pro Lys
                565                 570                 575
Gln Gln Gly Leu Arg Gly Thr Phe Gly Gly Asn Ile Ile Gln His Ser
            580                 585                 590
Ile Pro Ala Val Glu Leu Arg Gln Pro Phe Phe Pro Thr His Met Gly
        595                 600                 605
Pro Ile Lys Leu Arg Gln Phe His Arg Pro Pro Leu Lys Lys Tyr Ser
    610                 615                 620
Phe Gly Ala Leu Ser Gln Pro Gly Pro His Ser Val Gln Pro Leu Leu
625                 630                 635                 640
Lys His Ile Lys Lys Lys Ala Lys Met Arg Glu Gln Glu Arg Gln Ala
                645                 650                 655
Ser Gly Gly Gly Glu Met Phe Phe Met Arg Thr Pro Gln Asp Leu Thr
            660                 665                 670
Gly Lys Asp Gly Asp Leu Ile Leu Ala Glu Tyr Ser Glu Glu Asn Gly
        675                 680                 685
Pro Leu Met Met Gln Val Gly Met Ala Thr Lys Ile Lys Asn Tyr Tyr
    690                 695                 700
Lys Arg Lys Pro Gly Lys Asp Pro Gly Ala Pro Asp Cys Lys Tyr Gly
705                 710                 715                 720
Glu Thr Val Tyr Cys His Thr Ser Pro Phe Leu Gly Ser Leu His Pro
                725                 730                 735
Gly Gln Leu Leu Gln Ala Phe Glu Asn Asn Leu Phe Arg Ala Pro Ile
            740                 745                 750
Tyr Leu His Lys Met Pro Glu Thr Asp Phe Leu Ile Ile Arg Thr Arg
        755                 760                 765
Gln Gly Tyr Tyr Ile Arg Glu Leu Val Asp Ile Phe Val Val Gly Gln
    770                 775                 780
Gln Cys Pro Leu Phe Glu Val Pro Gly Pro Asn Ser Lys Arg Ala Asn
785                 790                 795                 800
Thr His Ile Arg Asp Phe Leu Gln Val Phe Ile Tyr Arg Leu Phe Trp
                805                 810                 815
```

```
Lys Ser Lys Asp Arg Pro Arg Arg Ile Arg Met Glu Asp Ile Lys Lys
                820                 825                 830
Ala Phe Pro Ser His Ser Glu Ser Ser Ile Arg Lys Arg Leu Lys Leu
        835                 840                 845
Cys Ala Asp Phe Lys Arg Thr Gly Met Asp Ser Asn Trp Trp Val Leu
    850                 855                 860
Lys Ser Asp Phe Arg Leu Pro Thr Glu Glu Glu Ile Arg Ala Met Val
865                 870                 875                 880
Ser Pro Glu Gln Cys Cys Ala Tyr Tyr Ser Met Ile Ala Ala Glu Gln
                885                 890                 895
Arg Leu Lys Asp Ala Gly Tyr Gly Glu Lys Ser Phe Phe Ala Pro Glu
            900                 905                 910
Glu Glu Asn Glu Glu Asp Phe Gln Met Lys Ile Asp Asp Glu Val Arg
        915                 920                 925
Thr Ala Pro Trp Asn Thr Thr Arg Ala Phe Ile Ala Ala Met Lys Gly
    930                 935                 940
Lys Cys Leu Leu Glu Val Thr Gly Val Ala Asp Pro Thr Gly Cys Gly
945                 950                 955                 960
Glu Gly Phe Ser Tyr Val Lys Ile Pro Asn Lys Pro Thr Gln Gln Lys
                965                 970                 975
Asp Asp Lys Glu Pro Gln Pro Val Lys Lys Thr Val Thr Gly Thr Asp
            980                 985                 990
Ala Asp Leu Arg Arg Leu Ser Leu Lys Asn Ala Lys Gln Leu Leu Arg
        995                 1000                1005
Lys Phe Gly Val Pro Glu Glu Glu Ile Lys Lys Leu Ser Arg Trp Glu
    1010                1015                1020
Val Ile Asp Val Val Arg Thr Met Ser Thr Glu Gln Ala Arg Ser Gly
1025                1030                1035                104
Glu Gly Pro Met Ser Lys Phe Ala Arg Gly Ser Arg Phe Ser Val Ala
                1045                1050                1055
Glu His Gln Glu Arg Tyr Lys Glu Glu Cys Gln Arg Ile Phe Asp Leu
            1060                1065                1070
Gln Asn Lys Val Leu Ser Ser Thr Glu Val Leu Ser Thr Asp Thr Asp
        1075                1080                1085
Ser Ser Ser Ala Glu Asp Ser Asp Phe Glu Glu Met Gly Lys Asn Ile
    1090                1095                1100
Glu Asn Met Leu Gln Asn Lys Lys Thr Ser Ser Gln Leu Ser Arg Glu
1105                1110                1115                112
Arg Glu Glu Gln Glu Arg Lys Glu Leu Gln Arg Met Leu Leu Ala Ala
                1125                1130                1135
Gly Ser Ala Ala Ser Gly Asn Asn His Arg Asp Asp Asp Thr Ala Ser
            1140                1145                1150
Val Thr Ser Leu Asn Ser Ser Ala Thr Gly Arg Cys Leu Lys Ile Tyr
        1155                1160                1165
Arg Thr Phe Arg Asp Glu Glu Gly Lys Glu Tyr Val Arg Cys Glu Thr
    1170                1175                1180
Val Arg Lys Pro Ala Val Ile Asp Ala Tyr Val Arg Ile Arg Thr Thr
1185                1190                1195                120
Lys Asp Glu Glu Phe Ile Arg Lys Phe Ala Leu Phe Asp Glu Gln His
                1205                1210                1215
```

```
Arg Glu Glu Met Arg Lys Glu Arg Arg Arg Ile Gln Glu Gln Leu Arg
            1220                1225                1230
Arg Leu Lys Arg Asn Gln Glu Lys Glu Lys Leu Lys Gly Pro Pro Glu
        1235                1240                1245
Lys Lys Pro Lys Lys Met Lys Glu Arg Pro Asp Leu Lys Leu Lys Cys
    1250                1255                1260
Gly Ala Cys Gly Ala Ile Gly His Met Arg Thr Asn Lys Phe Cys Pro
1265                1270                1275                    128
Leu Tyr Tyr Gln Thr Asn Ala Pro Pro Ser Asn Pro Val Ala Met Thr
                1285                1290                1295
Glu Glu Gln Glu Glu Glu Leu Glu Lys Thr Val Ile His Asn Asp Asn
            1300                1305                1310
Glu Glu Leu Ile Lys Val Glu Gly Thr Lys Ile Val Leu Gly Lys Gln
        1315                1320                1325
Leu Ile Glu Ser Ala Asp Glu Val Arg Arg Lys Ser Leu Val Leu Lys
    1330                1335                1340
Phe Pro Lys Gln Gln Leu Pro Pro Lys Lys Lys Arg Arg Val Gly Thr
1345                1350                1355                    136
Thr Val His Cys Asp Tyr Leu Asn Arg Pro His Lys Ser Ile His Arg
                1365                1370                1375
Arg Arg Thr Asp Pro Met Val Thr Leu Ser Ser Ile Leu Glu Ser Ile
            1380                1385                1390
Ile Asn Asp Met Arg Asp Leu Pro Asn Thr Tyr Pro Phe His Thr Pro
        1395                1400                1405
Val Asn Ala Lys Val Val Lys Asp Tyr Tyr Lys Ile Ile Thr Arg Pro
    1410                1415                1420
Met Asp Leu Gln Thr Leu Arg Glu Asn Val Arg Lys Arg Leu Tyr Pro
1425                1430                1435                    144
Ser Arg Glu Glu Phe Arg Glu His Leu Glu Leu Ile Val Lys Asn Ser
                1445                1450                1455
Ala Thr Tyr Asn Gly Pro Lys His Ser Leu Thr Gln Ile Ser Gln Ser
            1460                1465                1470
Met Leu Asp Leu Cys Asp Glu Lys Leu Lys Glu Lys Glu Asp Lys Leu
        1475                1480                1485
Ala Arg Leu Glu Lys Ala Ile Asn Pro Leu Leu Asp Asp Asp Asp Gln
    1490                1495                1500
Val Ala Phe Ser Phe Ile Leu Asp Asn Ile Val Thr Gln Lys Met Met
1505                1510                1515                    152
Ala Val Pro Asp Ser Trp Pro Phe His His Pro Val Asn Lys Lys Phe
                1525                1530                1535
Val Pro Asp Tyr Tyr Lys Val Ile Val Asn Pro Met Asp Leu Glu Thr
            1540                1545                1550
Ile Arg Lys Asn Ile Ser Lys His Lys Tyr Gln Ser Arg Glu Ser Phe
        1555                1560                1565
Leu Asp Asp Val Asn Leu Ile Leu Ala Asn Ser Val Lys Tyr Asn Gly
    1570                1575                1580
Pro Glu Ser Gln Tyr Thr Lys Thr Ala Gln Glu Ile Val Asn Val Cys
1585                1590                1595                    160
Tyr Gln Thr Leu Thr Glu Tyr Asp Glu His Leu Thr Gln Leu Glu Lys
                1605                1610                1615
```

```
Asp Ile Cys Thr Ala Lys Glu Ala Ala Leu Glu Glu Ala Glu Leu Glu
            1620                1625                1630
Ser Leu Asp Pro Met Thr Pro Gly Pro Tyr Thr Pro Gln Pro Pro Asp
        1635                1640                1645
Leu Tyr Asp Thr Asn Thr Ser Leu Ser Met Ser Arg Asp Ala Ser Val
    1650                1655                1660
Phe Gln Asp Glu Ser Asn Met Ser Val Leu Asp Ile Pro Ser Ala Thr
1665                1670                1675                168
Pro Glu Lys Gln Val Thr Gln Glu Gly Glu Asp Gly Asp Gly Asp Leu
                1685                1690                1695
Ala Asp Glu Glu Glu Gly Thr Val Gln Gln Pro Gln Ala Ser Val Leu
            1700                1705                1710
Tyr Glu Asp Leu Leu Met Ser Glu Gly Glu Asp Asp Glu Glu Asp Ala
        1715                1720                1725
Gly Ser Asp Glu Glu Gly Asp Asn Pro Phe Ser Ala Ile Gln Leu Ser
    1730                1735                1740
Glu Ser Gly Ser Asp Ser Asp Val Gly Ser Gly Gly Ile Arg Pro Lys
1745                1750                1755                176
Gln Pro Arg Met Leu Gln Glu Asn Thr Arg Met Asp Met Glu Asn Glu
                1765                1770                1775
Glu Ser Met Met Ser Tyr Glu Gly Asp Gly Gly Glu Ala Ser His Gly
            1780                1785                1790
Leu Glu Asp Ser Asn Ile Ser Tyr Gly Ser Tyr Glu Glu Pro Asp Pro
        1795                1800                1805
Lys Ser Asn Thr Gln Asp Thr Ser Phe Ser Ser Ile Gly Gly Tyr Glu
    1810                1815                1820
Val Ser Glu Glu Glu Glu Asp Glu Glu Glu Glu Gln Arg Ser Gly
1825                1830                1835                184
Pro Ser Val Leu Ser Gln Val His Leu Ser Glu Asp Glu Glu Asp Ser
                1845                1850                1855
Glu Asp Phe His Ser Ile Ala Gly Asp Ser Asp Leu Asp Ser Asp Glu
            1860                1865                1870
```

<210> 7
<211> 429
<212> PRT
<213> C. albicans

<400> 7

```
Glu Leu Leu Leu Asn Asn Pro Leu Asp Asn Ser Lys Gln Asn Arg Gln
 1               5                  10                  15
Lys Ile Glu Asn Asp Asn Thr Thr Asn Asn Tyr Asn Gln Asn Asn Ser
            20                  25                  30
Asn Val Gln Asp Glu Glu Glu Asp Asp Asp Ile Phe Asn Gly Gln Ile
        35                  40                  45
Asn Leu Asp Lys Leu Lys Leu Asp Met Asn Asp Pro Asn Leu Leu Phe
    50                  55                  60
Val Pro Ser Lys Lys Val Asp Ala Thr Lys Ser Val Val Pro Ser Thr
65                  70                  75                  80
Asp Lys Leu Leu Glu Leu Lys Phe Asn Ile Ser Asn Asp Gln Glu Tyr
```

```
                85                  90                  95
Glu Leu Leu Arg Lys Asn Tyr Asn Thr Lys Gln Arg Ser Gln Leu Ser
            100                 105                 110
Asn Leu Asn Ile Glu His Ser Val Pro Ala Leu Arg Leu Gln Thr Pro
        115                 120                 125
Tyr Tyr Lys Val Lys Leu Ser Thr Asp Glu Thr Arg Ser Phe His Arg
    130                 135                 140
Pro Val Phe Asn Val Arg Pro Gly Thr Leu Val Ser Phe Ser Lys Leu
145                 150                 155                 160
Lys Leu Arg Lys Arg Lys Lys Asp Lys Gly Lys Ser Leu Gln Gln Ile
                165                 170                 175
Phe Ser Lys Thr Ser Asp Leu Thr Val Ala Asp Thr Gly Asn Ile Ile
            180                 185                 190
Ala Leu Glu Tyr Ser Glu Gln Tyr Pro Pro Ile Leu Ser Asn Phe Gly
        195                 200                 205
Met Gly Ser Lys Leu Ile Asn Tyr Tyr Arg Lys Glu Arg Pro Asn Asp
    210                 215                 220
Thr Ser Arg Pro Lys Ala Gln Ile Gly Glu Thr His Ile Leu Gly Val
225                 230                 235                 240
Glu Asp Arg Ser Pro Phe Trp Asn Phe Gly Glu Val Ala Pro Gly Asp
                245                 250                 255
Phe Val Pro Thr Leu Tyr Asn Asn Met Val Arg Ala Pro Ile Phe Lys
            260                 265                 270
His Asp Asn Lys Pro Thr Asp Phe Leu Leu Val Lys Ser Gln Gly Ala
        275                 280                 285
Gly Ser His Gln Lys Phe Tyr Leu Arg Gly Ile Asn Phe Asn Phe Ala
    290                 295                 300
Val Gly Asn Thr Phe Pro Val Glu Val Pro Ala Pro His Ser Arg Lys
305                 310                 315                 320
Val Thr Asn Ile Ser Lys Asn Arg Leu Lys Met Val Val Phe Arg Val
                325                 330                 335
Met Asn Ser Leu Gly Val Pro Arg Ile Ser Val Lys Asp Val Ser Lys
            340                 345                 350
His Phe Pro Glu His Ser Asp Met Gln Asn Arg Gln Arg Leu Lys Glu
        355                 360                 365
Phe Met Glu Tyr Gln Arg Gln Gly Glu Asp Gln Gly Tyr Trp Lys Val
    370                 375                 380
Arg Gly Leu Asn Asp Val Ile Pro Gly Glu Glu Glu Ile Arg Thr Met
385                 390                 395                 400
Ile Thr Pro Glu Asp Ser Ser Leu Met Asp Thr Met Gln Phe Gly Gln
                405                 410                 415
Gln Val Leu Asp Asp Asn Met Val Leu Phe Gly Glu Gln
            420                 425
```

<210> 8
<211> 434
<212> PRT
<213> S. cerevisiae

<400> 8

```
Met Thr Pro Asn Leu Lys Phe Ser Gly Gly Tyr Lys Leu Lys Ser Leu
 1               5                  10                  15
Ile Glu Asp Val Ala Glu Asp Trp Gln Trp Asp Glu Asp Met Ile Ile
            20                  25                  30
Asp Ala Lys Leu Lys Glu Ser Lys His Ala Glu Leu Asn Met Asn Asp
        35                  40                  45
Glu Lys Leu Leu Leu Met Ile Glu Lys Thr Asn Asn Leu Ala Gln Gln
    50                  55                  60
Lys Gln Gln Leu Asp Ser Ser Asn Leu Ile Leu Pro Leu Asn Glu Thr
65                  70                  75                  80
Ile Leu Gln Gln Lys Phe Asn Leu Ser Asn Asp Asp Lys Tyr Gln Ile
                85                  90                  95
Leu Lys Lys Thr His Gln Thr Lys Val Arg Ser Thr Ile Ser Asn Leu
            100                 105                 110
Asn Ile Gln His Ser Gln Pro Ala Ile Asn Leu Gln Ser Pro Phe Tyr
        115                 120                 125
Lys Val Ala Val Pro Arg Tyr Gln Leu Arg His Phe His Arg Glu Asn
    130                 135                 140
Phe Gly Ser His Ile Arg Pro Gly Thr Lys Ile Val Phe Ser Lys Leu
145                 150                 155                 160
Lys Ala Arg Lys Arg Lys Arg Asp Lys Gly Lys Asp Val Lys Glu Ser
                165                 170                 175
Phe Ser Thr Ser Gln Asp Leu Thr Ile Gly Asp Thr Ala Pro Val Tyr
            180                 185                 190
Leu Met Glu Tyr Ser Glu Gln Thr Pro Val Ala Leu Ser Lys Phe Gly
        195                 200                 205
Met Ala Asn Lys Leu Ile Asn Tyr Tyr Arg Lys Ala Asn Glu Gln Asp
    210                 215                 220
Thr Leu Arg Pro Lys Leu Pro Val Gly Glu Thr His Val Leu Gly Val
225                 230                 235                 240
Gln Asp Lys Ser Pro Phe Trp Asn Phe Gly Phe Val Glu Pro Gly His
                245                 250                 255
Ile Val Pro Thr Leu Tyr Asn Asn Met Ile Arg Ala Pro Val Phe Lys
            260                 265                 270
His Asp Ile Ser Gly Thr Asp Phe Leu Leu Thr Lys Ser Ser Gly Phe
        275                 280                 285
Gly Ile Ser Asn Arg Phe Tyr Leu Arg Asn Ile Asn His Leu Phe Thr
    290                 295                 300
Val Gly Gln Thr Phe Pro Val Glu Glu Ile Pro Gly Pro Asn Ser Arg
305                 310                 315                 320
Lys Val Thr Ser Met Lys Ala Thr Arg Leu Lys Met Ile Ile Tyr Arg
                325                 330                 335
Ile Leu Asn His Asn His Ser Lys Ala Ile Ser Ile Asp Pro Ile Ala
            340                 345                 350
Lys His Phe Pro Asp Gln Asp Tyr Gly Gln Asn Arg Gln Lys Val Lys
        355                 360                 365
Glu Phe Met Lys Tyr Gln Arg Asp Gly Pro Glu Lys Gly Leu Trp Arg
    370                 375                 380
Leu Lys Asp Asp Glu Lys Leu Leu Asp Asn Glu Ala Val Lys Ser Leu
385                 390                 395                 400
```

```
Ile Thr Pro Glu Gln Ile Ser Gln Val Glu Ser Met Ser Gln Gly Leu
                405                 410                 415
Gln Phe Gln Glu Asp Asn Glu Ala Tyr Asn Phe Asp Ser Lys Leu Lys
            420                 425                 430
Ser Leu
```

<210> 9
<211> 415
<212> PRT
<213> S. pombe

<400> 9

```
Val Asn Lys Thr Asn Gln Ser Ser Phe Phe Ile Asp Lys Ser Leu Val
 1               5                  10                  15
Asp Ile Asp Phe Ala Phe Asp Glu Asn Ile Phe Asp Gly Asp Thr Gly
            20                  25                  30
Thr Ser Lys Val Val Leu Asn Leu Asn Asp Pro Lys Leu Leu Leu Gln
        35                  40                  45
Pro Gln Leu Pro Lys Lys Glu Asp Ser Gln Arg Ser Phe Ala Asp Thr
    50                  55                  60
His Gln Arg Asn Ser Leu Ala Trp Lys Phe Asn Ile Ser Asn Asp Pro
65                  70                  75                  80
Ala Tyr Glu Met Leu Lys Gln Asn His Gln Ser Lys Val Arg Asn Thr
                85                  90                  95
Leu Ser Gln Leu Ala Ile Glu His Ala Ala Phe Ala Glu Lys Leu Thr
            100                 105                 110
Phe Pro Tyr Tyr Lys Thr Arg Leu Ser Lys Arg Ala Val Arg Ser Tyr
        115                 120                 125
His Arg Pro Thr Met Ser Phe Lys Pro Asn Ala Ala Ile Val Phe Ser
    130                 135                 140
Pro Leu Ile Val Arg Lys Arg Ser Lys Asp Lys His Lys Ser Glu Arg
145                 150                 155                 160
Glu Leu Ile Pro Thr Thr Lys Glu Ile Thr Met Gly Asp Thr Thr His
                165                 170                 175
Ala Ile Leu Val Glu Phe Ser Glu Glu His Pro Ala Val Leu Ser Asn
            180                 185                 190
Ala Gly Met Ala Ser Arg Ile Val Asn Tyr Tyr Arg Lys Lys Asn Glu
        195                 200                 205
Gln Asp Glu Ser Arg Pro Lys Leu Glu Val Gly Glu Ser His Val Leu
    210                 215                 220
Asp Val Gln Asp Arg Ser Pro Phe Trp Asn Phe Gly Ser Val Glu Pro
225                 230                 235                 240
Gly Glu Ile Thr Pro Thr Leu Tyr Asn Lys Met Ile Arg Ala Pro Leu
                245                 250                 255
Phe Lys His Glu Val Pro Pro Thr Asp Phe Ile Leu Ile Arg Asn Ser
            260                 265                 270
Ser Ser Tyr Gly Ser Lys Tyr Tyr Leu Lys Asn Ile Asn His Met Phe
        275                 280                 285
Val Ser Gly Gln Thr Phe Pro Val Thr Asp Val Pro Gly Pro His Ser
```

```
    290                 295                 300
Arg Lys Val Thr Thr Ala Ser Lys Asn Arg Leu Lys Met Leu Val Phe
305                 310                 315                 320
Arg Leu Ile Arg Arg Ser Pro Asn Gly Gly Leu Phe Ile Arg Gln Leu
                325                 330                 335
Ser Lys His Phe Ser Asp Gln Asn Glu Met Gln Ile Arg Gln Arg Leu
            340                 345                 350
Lys Glu Phe Met Glu Tyr Lys Lys Lys Gly Asp Gly Pro Gly Tyr Trp
        355                 360                 365
Lys Leu Lys Ser Asn Glu Val Val Pro Asp Glu Ala Gly Thr Arg Ser
    370                 375                 380
Met Val Ser Pro Glu Thr Val Cys Leu Leu Glu Ser Met Gln Val Gly
385                 390                 395                 400
Val Arg Gln Leu Glu Asp Ala Gly Tyr Gly Lys Thr Met Asp Glu
                405                 410                 415

      <210> 10
      <211> 481
      <212> PRT
      <213> Human

      <400> 10
Ser Leu Ala Gly Trp Leu Pro Ser Ser Met Thr Arg Asn Ala Met Ala
 1               5                  10                  15
Tyr Asn Val Gln Gln Gly Phe Ala Ala Thr Leu Asp Asp Asp Lys Pro
            20                  25                  30
Trp Tyr Ser Ile Phe Pro Ile Asp Asn Glu Asp Leu Val Tyr Gly Arg
        35                  40                  45
Trp Glu Asp Asn Ile Ile Trp Asp Ala Gln Ala Met Pro Arg Leu Leu
    50                  55                  60
Glu Pro Pro Val Leu Thr Leu Asp Pro Asn Asp Glu Asn Leu Ile Leu
65                  70                  75                  80
Glu Ile Pro Asp Glu Lys Glu Glu Ala Thr Ser Asn Ser Pro Ser Lys
                85                  90                  95
Glu Ser Lys Lys Glu Ser Ser Leu Lys Lys Ser Arg Ile Leu Leu Gly
            100                 105                 110
Lys Thr Gly Val Ile Lys Glu Glu Pro Gln Gln Asn Met Ser Gln Pro
        115                 120                 125
Glu Val Lys Asp Pro Trp Asn Leu Ser Asn Asp Glu Tyr Tyr Tyr Pro
    130                 135                 140
Lys Gln Gln Gly Leu Arg Gly Thr Phe Gly Gly Asn Ile Ile Gln His
145                 150                 155                 160
Ser Ile Pro Ala Val Glu Leu Arg Gln Pro Phe Phe Pro Thr His Met
                165                 170                 175
Gly Pro Ile Lys Leu Arg Gln Phe His Arg Pro Pro Leu Lys Lys Tyr
            180                 185                 190
Ser Phe Gly Ala Leu Ser Gln Pro Gly Pro His Ser Val Gln Pro Leu
        195                 200                 205
Leu Lys His Ile Lys Lys Lys Ala Lys Met Arg Glu Gln Glu Arg Gln
    210                 215                 220
```

```
Ala Ser Gly Gly Gly Glu Met Phe Phe Met Arg Thr Pro Gln Asp Leu
225                 230                 235                 240
Thr Gly Lys Asp Gly Asp Leu Ile Leu Ala Glu Tyr Ser Glu Glu Asn
                245                 250                 255
Gly Pro Leu Met Met Gln Val Gly Met Ala Thr Lys Ile Lys Asn Tyr
            260                 265                 270
Tyr Lys Arg Lys Pro Gly Lys Asp Pro Gly Ala Pro Asp Cys Lys Tyr
        275                 280                 285
Gly Glu Thr Val Tyr Cys His Thr Ser Pro Phe Leu Gly Ser Leu His
    290                 295                 300
Pro Gly Gln Leu Leu Gln Ala Phe Glu Asn Asn Leu Phe Arg Ala Pro
305                 310                 315                 320
Ile Tyr Leu His Lys Met Pro Glu Thr Asp Phe Leu Ile Ile Arg Thr
                325                 330                 335
Arg Gln Gly Tyr Tyr Ile Arg Glu Leu Val Asp Ile Phe Val Val Gly
            340                 345                 350
Gln Gln Cys Pro Leu Phe Glu Val Pro Gly Pro Asn Ser Lys Arg Ala
        355                 360                 365
Asn Thr His Ile Arg Asp Phe Leu Gln Val Phe Ile Tyr Arg Leu Phe
    370                 375                 380
Trp Lys Ser Lys Asp Arg Pro Arg Arg Ile Arg Met Glu Asp Ile Lys
385                 390                 395                 400
Lys Ala Phe Pro Ser His Ser Glu Ser Ser Ile Arg Lys Arg Leu Lys
                405                 410                 415
Leu Cys Ala Asp Phe Lys Arg Thr Gly Met Asp Ser Asn Trp Trp Val
            420                 425                 430
Leu Lys Ser Asp Phe Arg Leu Pro Thr Glu Glu Glu Ile Arg Ala Met
        435                 440                 445
Val Ser Pro Glu Gln Cys Cys Ala Tyr Tyr Ser Met Ile Ala Ala Glu
    450                 455                 460
Gln Arg Leu Lys Asp Ala Gly Tyr Gly Glu Lys Ser Phe Phe Ala Pro
465                 470                 475                 480
Glu
```

<210> 11
<211> 74
<212> PRT
<213> C. albicans

<400> 11

```
Asp Ala Glu Asn Gly Asp Asp Ile Asn Lys Asp Lys Glu Lys Glu Val
 1               5                  10                  15
Glu Lys Glu Lys Glu Gln Glu Arg Glu Glu Glu Lys Gly Lys Asp Lys
            20                  25                  30
Glu Lys Asp Lys Asp Lys Glu Lys Asp Lys Thr Glu Lys Glu Lys Ser
        35                  40                  45
Lys Lys Ser Lys Glu Gln Asp Thr Glu Ile Asp Val Glu Glu Glu Leu
    50                  55                  60
Ala Pro Trp Asn Leu Ser Arg Asn Phe Val
```

65 70

<210> 12
<211> 18
<212> DNA
<213> "Artificial Sequence"

<220>

<223> Inosine

<400> 12
ccwggwccwa aytcnadd 18

<210> 13
<211> 23
<212> DNA
<213> "Artificial Sequence"

<400> 13
gayccwachg gwtgtggwga agg 23

<210> 14
<211> 24
<212> DNA
<213> "Artificial Sequence"

<400> 14
cctttcwcca cawccagtwg grtc 24

<210> 15
<211> 19
<212> DNA
<213> "Artificial Sequence"

<220>

<223> Inosine

<400> 15
ttrtthcayc tnartgwcc 19

<210> 16
<211> 30
<212> DNA
<213> "Artificial Sequence"

<400> 16
ccgctcgaga tgacacccaa cttaaagttc 30

<210> 17
<211> 29
<212> DNA
<213> "Artificial Sequence"

<400> 17
cgcggatcca gagattttag cttagaatc 29

<210> 18
<211> 37
<212> DNA
<213> "Artificial Sequence"

<400> 18
ggaattccat atgcttttgc tcaacaatcc cttggac 37

<210> 19
<211> 32
<212> DNA
<213> "Artificial Sequence"

<400> 19
cgcggatccc tgctctgctc accgaataac ac 32

<210> 20
<211> 37
<212> DNA
<213> "Artificial Sequence"

<400> 20
ggaattccat atgagcctgg caggctggct tccttct 37

<210> 21
<211> 33
<212> DNA
<213> "Artificial Sequence"

<400> 21
ccgctcgagt tctggagcaa aaaaggattt ctc 33

<210> 22
<211> 0
<212> DNA
<213> Human

<400> 22
Met Gly Pro Gly Cys Asp Leu Leu Leu Arg Thr Ala Ala Thr Ile Thr
1 5 10 15
Ala Ala Ala Ile Met Ser Asp Thr Asp Ser Asp Glu Asp Ser Ala Gly
20 25 30

```
Gly Gly Pro Phe Ser Leu Ala Gly Phe Leu Phe Gly Asn Ile Asn Gly
        35                  40                  45
Ala Gly Gln Leu Glu Gly Glu Ser Val Leu Asp Asp Glu Cys Lys Lys
    50                  55                  60
His Leu Ala Gly Leu Gly Ala Leu Gly Leu Gly Ser Leu Ile Thr Glu
65                  70                  75                  80
Leu Thr Ala Asn Glu Glu Leu Thr Gly Thr Asp Gly Ala Leu Val Asn
                85                  90                  95
Asp Glu Gly Trp Val Arg Ser Thr Glu Asp Ala Val Asp Tyr Ser Asp
            100                 105                 110
Ile Asn Glu Val Ala Glu Asp Glu Ser Arg Arg Tyr Gln Gln Thr Met
        115                 120                 125
Gly Ser Leu Gln Pro Leu Cys His Ser Asp Tyr Asp Glu Asp Asp Tyr
    130                 135                 140
Asp Ala Asp Cys Glu Asp Ile Asp Cys Lys Leu Met Pro Pro Pro Pro
145                 150                 155                 160
Pro Pro Pro Gly Pro Met Lys Lys Asp Lys Asp Gln Asp Ser Ile Thr
                165                 170                 175
Gly Glu Lys Val Asp Phe Ser Ser Ser Ser Asp Ser Glu Ser Glu Met
            180                 185                 190
Gly Pro Gln Glu Ala Thr Gln Ala Glu Ser Glu Asp Gly Lys Leu Thr
        195                 200                 205
Leu Pro Leu Ala Gly Ile Met Gln His Asp Ala Thr Lys Leu Leu Pro
    210                 215                 220
Ser Val Thr Glu Leu Phe Pro Glu Phe Arg Pro Gly Lys Val Leu Arg
225                 230                 235                 240
Phe Leu Arg Leu Phe Gly Pro Gly Lys Asn Val Pro Ser Val Trp Arg
                245                 250                 255
Ser Ala Arg Arg Lys Arg Lys Lys Lys His Arg Glu Leu Ile Gln Glu
            260                 265                 270
Glu Gln Ile Gln Glu Val Glu Cys Ser Val Glu Ser Glu Val Ser Gln
        275                 280                 285
Lys Ser Leu Trp Asn Tyr Asp Tyr Ala Pro Pro Pro Pro Pro Glu Gln
    290                 295                 300
Cys Leu Ser Asp Asp Glu Ile Thr Met Met Ala Pro Val Glu Ser Lys
305                 310                 315                 320
Phe Ser Gln Ser Thr Gly Asp Ile Asp Lys Val Thr Asp Thr Lys Pro
                325                 330                 335
Arg Val Ala Glu Trp Arg Tyr Gly Pro Ala Arg Leu Trp Tyr Asp Met
            340                 345                 350
Leu Gly Val Pro Glu Asp Gly Ser Gly Phe Asp Tyr Gly Phe Lys Leu
        355                 360                 365
Arg Lys Thr Glu His Glu Pro Val Ile Lys Ser Arg Met Ile Glu Glu
    370                 375                 380
Phe Arg Lys Leu Glu Glu Asn Asn Gly Thr Asp Leu Leu Ala Asp Glu
385                 390                 395                 400
Asn Phe Leu Met Val Thr Gln Leu His Trp Glu Asp Asp Ile Ile Trp
                405                 410                 415
Asp Gly Glu Asp Val Lys His Lys Gly Thr Lys Pro Gln Arg Ala Ser
            420                 425                 430
```

```
Leu Ala Gly Trp Leu Pro Ser Ser Met Thr Arg Asn Ala Met Ala Tyr
        435                 440                 445
Asn Val Gln Gln Gly Phe Ala Ala Thr Leu Asp Asp Asp Lys Pro Trp
    450                 455                 460
Tyr Ser Ile Phe Pro Ile Asp Asn Glu Asp Leu Val Tyr Gly Arg Trp
465                 470                 475                 480
Glu Asp Asn Ile Ile Trp Asp Ala Gln Ala Met Pro Arg Leu Leu Glu
                485                 490                 495
Pro Pro Val Leu Thr Leu Asp Pro Asn Asp Glu Asn Leu Ile Leu Glu
            500                 505                 510
Ile Pro Asp Glu Lys Glu Glu Ala Thr Ser Asn Ser Pro Ser Lys Glu
        515                 520                 525
Ser Lys Lys Glu Ser Ser Leu Lys Lys Ser Arg Ile Leu Leu Gly Lys
    530                 535                 540
Thr Gly Val Ile Lys Glu Glu Pro Gln Gln Asn Met Ser Gln Pro Glu
545                 550                 555                 560
Val Lys Asp Pro Trp Asn Leu Ser Asn Asp Glu Tyr Tyr Tyr Pro Lys
                565                 570                 575
Gln Gln Gly Leu Arg Gly Thr Phe Gly Gly Asn Ile Ile Gln His Ser
            580                 585                 590
Ile Pro Ala Val Glu Leu Arg Gln Pro Phe Phe Pro Thr His Met Gly
        595                 600                 605
Pro Ile Lys Leu Arg Gln Phe His Arg Pro Pro Leu Lys Lys Tyr Ser
    610                 615                 620
Phe Gly Ala Leu Ser Gln Pro Gly Pro His Ser Val Gln Pro Leu Leu
625                 630                 635                 640
Lys His Ile Lys Lys Lys Ala Lys Met Arg Glu Gln Glu Arg Gln Ala
                645                 650                 655
Ser Gly Gly Gly Glu Met Phe Phe Met Arg Thr Pro Gln Asp Leu Thr
            660                 665                 670
Gly Lys Asp Gly Asp Leu Ile Leu Ala Glu Tyr Ser Glu Glu Asn Gly
        675                 680                 685
Pro Leu Met Met Gln Val Gly Met Ala Thr Lys Ile Lys Asn Tyr Tyr
    690                 695                 700
Lys Arg Lys Pro Gly Lys Asp Pro Gly Ala Pro Asp Cys Lys Tyr Gly
705                 710                 715                 720
Glu Thr Val Tyr Cys His Thr Ser Pro Phe Leu Gly Ser Leu His Pro
                725                 730                 735
Gly Gln Leu Leu Gln Ala Phe Glu Asn Asn Leu Phe Arg Ala Pro Ile
            740                 745                 750
Tyr Leu His Lys Met Pro Glu Thr Asp Phe Leu Ile Ile Arg Thr Arg
        755                 760                 765
Gln Gly Tyr Tyr Ile Arg Glu Leu Val Asp Ile Phe Val Val Gly Gln
    770                 775                 780
Gln Cys Pro Leu Phe Glu Val Pro Gly Pro Asn Ser Lys Arg Ala Asn
785                 790                 795                 800
Thr His Ile Arg Asp Phe Leu Gln Val Phe Ile Tyr Arg Leu Phe Trp
                805                 810                 815
Lys Ser Lys Asp Arg Pro Arg Arg Ile Arg Met Glu Asp Ile Lys Lys
            820                 825                 830
```

```
Ala Phe Pro Ser His Ser Glu Ser Ser Ile Arg Lys Arg Leu Lys Leu
        835                 840                 845
Cys Ala Asp Phe Lys Arg Thr Gly Met Asp Ser Asn Trp Trp Val Leu
    850                 855                 860
Lys Ser Asp Phe Arg Leu Pro Thr Glu Glu Glu Ile Arg Ala Met Val
865                 870                 875                 880
Ser Pro Glu Gln Cys Cys Ala Tyr Tyr Ser Met Ile Ala Ala Glu Gln
                885                 890                 895
Arg Leu Lys Asp Ala Gly Tyr Gly Glu Lys Ser Phe Phe Ala Pro Glu
            900                 905                 910
Glu Glu Asn Glu Glu Asp Phe Gln Met Lys Ile Asp Asp Glu Val Arg
        915                 920                 925
Thr Ala Pro Trp Asn Thr Thr Arg Ala Phe Ile Ala Ala Met Lys Gly
    930                 935                 940
Lys Cys Leu Leu Glu Val Thr Gly Val Ala Asp Pro Thr Gly Cys Gly
945                 950                 955                 960
Glu Gly Phe Ser Tyr Val Lys Ile Pro Asn Lys Pro Thr Gln Gln Lys
                965                 970                 975
Asp Asp Lys Glu Pro Gln Pro Val Lys Lys Thr Val Thr Gly Thr Asp
            980                 985                 990
Ala Asp Leu Arg Arg Leu Ser Leu Lys Asn Ala Lys Gln Leu Leu Arg
        995                 1000                1005
Lys Phe Gly Val Pro Glu Glu Glu Ile Lys Lys Leu Ser Arg Trp Glu
    1010                1015                1020
Val Ile Asp Val Val Arg Thr Met Ser Thr Glu Gln Ala Arg Ser Gly
1025                1030                1035                104
Glu Gly Pro Met Ser Lys Phe Ala Arg Gly Ser Arg Phe Ser Val Ala
                1045                1050                1055
Glu His Gln Glu Arg Tyr Lys Glu Glu Cys Gln Arg Ile Phe Asp Leu
            1060                1065                1070
Gln Asn Lys Val Leu Ser Ser Thr Glu Val Leu Ser Thr Asp Thr Asp
        1075                1080                1085
Ser Ser Ser Ala Glu Asp Ser Asp Phe Glu Glu Met Gly Lys Asn Ile
    1090                1095                1100
Glu Asn Met Leu Gln Asn Lys Lys Thr Ser Ser Gln Leu Ser Arg Glu
1105                1110                1115                112
Arg Glu Glu Gln Glu Arg Lys Glu Leu Gln Arg Met Leu Leu Ala Ala
                1125                1130                1135
Gly Ser Ala Ala Ser Gly Asn Asn His Arg Asp Asp Asp Thr Ala Ser
            1140                1145                1150
Val Thr Ser Leu Asn Ser Ser Ala Thr Gly Arg Cys Leu Lys Ile Tyr
        1155                1160                1165
Arg Thr Phe Arg Asp Glu Glu Gly Lys Glu Tyr Val Arg Cys Glu Thr
    1170                1175                1180
Val Arg Lys Pro Ala Val Ile Asp Ala Tyr Val Arg Ile Arg Thr Thr
1185                1190                1195                120
Lys Asp Glu Glu Phe Ile Arg Lys Phe Ala Leu Phe Asp Glu Gln His
                1205                1210                1215
Arg Glu Glu Met Arg Lys Glu Arg Arg Arg Ile Gln Glu Gln Leu Arg
            1220                1225                1230
```

Arg Leu

<210> 23
<211> 52
<212> DNA
<213> C. albicans

<400> 23
acgcgtcgac atgcatcatc atcatcatca tatggaggat ctacccaggg at 52

<210> 24
<211> 38
<212> DNA
<213> C. albicans

<400> 24
atagttagcg gccgcacact gctggtgtca accaacaa 38

<210> 25
<211> 52
<212> DNA
<213> Human

<400> 25
acgcgtcgac atgcatcatc atcatcatca tatgggaccc ggctgcgatt tg 52

<210> 26
<211> 31
<212> DNA
<213> Human

<400> 26
gttgctctgc agctatcatg ctataataag c 31

<210> 27
<211> 31
<212> DNA
<213> Human

<400> 27
tgatagctgc agagcaacga ctgaaggatg c 31

<210> 28
<211> 32
<212> DNA
<213> Human

<400> 28
ccggtaccct cccgatgttg ttcatcaaaa ag 32

<210> 29
<211> 30
<212> DNA
<213> "Artificial Sequence"

<400> 29
acgcgtcgac atccaagttc aagttgtctg 30

<210> 30
<211> 42
<212> DNA
<213> "Artificial Sequence"

<400> 30
cgcggatccg cgctgcagtt ttcacatctt cttcttctgc ca 42

<210> 31
<211> 41
<212> DNA
<213> "Artificial Sequence"

<400> 31
aaaactgcag cgcggatccg cgtgcaggtg acgttattgg a 41

<210> 32
<211> 36
<212> DNA
<213> "Artificial Sequence"

<400> 32
atagtttagc ggccgccttg tgacaagaag tgacac 36